From "Genome: The Autobiography of a Species in 23 Chapters"

CHROMOSOME 4 By Matt Ridley

Fate

> Sir, what ye're telling us is nothing but scientific Calvinism.
>
> *Anonymous Scottish soldier to William Bateson after a popular lecture*[1]

Open any catalogue of the human genome and you will be confronted not with a list of human potentialities, but a list of diseases, mostly ones named after pairs of obscure central-European doctors. This gene causes Niemann–Pick disease; that one causes Wolf–Hirschhorn syndrome. The impression given is that genes are there to cause diseases. 'New gene for mental illness', announces a website on genes that reports the latest news from the front, 'The gene for early-onset dystonia. Gene for kidney cancer isolated. Autism linked to serotonin transporter gene. A new Alzheimer's gene. The genetics of obsessive behaviour.'

Yet to define genes by the diseases they cause is about as absurd as defining organs of the body by the diseases they get: livers are there to cause cirrhosis, hearts to cause heart attacks and brains to cause strokes. It is a measure, not of our knowledge but of our ignorance that this is the way the genome catalogues read. It is

literally true that the only thing we know about some genes is that their malfunction causes a particular disease. This is a pitifully small thing to know about a gene, and a terribly misleading one. It leads to the dangerous shorthand that runs as follows: 'X has got the Wolf–Hirschhorn gene.' Wrong. We all have the Wolf–Hirschhorn gene, except, ironically, people who have Wolf–Hirschhorn syndrome. Their sickness is caused by the fact that the gene is missing altogether. In the rest of us, the gene is a positive, not a negative force. The sufferers have the mutation, not the gene.

Wolf–Hirschhorn syndrome is so rare and so serious – its gene is so vital – that its victims die young. Yet the gene, which lies on chromosome 4, is actually the most famous of all the 'disease' genes because of a very different disease associated with it: Huntington's chorea. A mutated version of the gene causes Huntington's chorea; a complete lack of the gene causes Wolf–Hirschhorn syndrome. We know very little about what the gene is there to do in everyday life, but we now know in excruciating detail how and why and where it can go wrong and what the consequence for the body is. The gene contains a single 'word', repeated over and over again: CAG, CAG, CAG, CAG . . . The repetition continues sometimes just six times, sometimes thirty, sometimes more than a hundred times. Your destiny, your sanity and your life hang by the thread of this repetition. If the 'word' is repeated thirty-five times or fewer, you will be fine. Most of us have about ten to fifteen repeats. If the 'word' is repeated thirty-nine times or more, you will in mid-life slowly start to lose your balance, grow steadily more incapable of looking after yourself and die prematurely. The decline begins with a slight deterioration of the intellectual faculties, is followed by jerking limbs and descends into deep depression, occasional hallucination and delusions. There is no appeal: the disease is incurable. But it takes between fifteen and twenty-five horrifying years to run its course. There are few worse fates. Indeed, many of the early psychological symptoms of the disease are just as bad in those who live in an affected family but do not get the disease: the strain and stress of waiting for it to strike are devastating.

The cause is in the genes and nowhere else. Either you have the Huntington's mutation and will get the disease or not. This is determinism, predestination and fate on a scale of which Calvin never dreamed. It seems at first sight to be the ultimate proof that the genes are in charge and that there is nothing we can do about it. It does not matter if you smoke, or take vitamin pills, if you work out or become a couch potato. The age at which the madness will appear depends strictly and implacably on the number of repetitions of the 'word' CAG in one place in one gene. If you have thirty-nine, you have a ninety per cent probability of dementia by the age of seventy-five and will on average get the first symptoms at sixty-six; if forty, on average you will succumb at fifty-nine; if forty-one, at fifty-four; if forty-two, at thirty-seven; and so on until those who have fifty repetitions of the 'word' will lose their minds at roughly twenty-seven years of age. The scale is this: if your chromosomes were long enough to stretch around the equator, the difference between health and insanity would be less than one extra inch.[2]

No horoscope matches this accuracy. No theory of human causality, Freudian, Marxist, Christian or animist, has ever been so precise. No prophet in the Old Testament, no entrail-gazing oracle in ancient Greece, no crystal-ball gipsy clairvoyant on the pier at Bognor Regis ever pretended to tell people exactly when their lives would fall apart, let alone got it right. We are dealing here with a prophecy of terrifying, cruel and inflexible truth. There are a billion three-letter 'words' in your genome. Yet the length of just this one little motif is all that stands between each of us and mental illness.

Huntington's disease, which became notorious when it killed the folk singer Woody Guthrie in 1967, was first diagnosed by a doctor, George Huntington, in 1872 on the eastern tip of Long Island. He noticed that it seemed to run in families. Later work revealed that the Long Island cases were part of a much larger family tree originating in New England. In twelve generations of this pedigree more than a thousand cases of the disease could be found. All were descended from two brothers who emigrated from Suffolk in 1630. Several of their descendants were burnt as witches in Salem in 1693,

possibly because of the alarming nature of the disease. But because the mutation only makes itself manifest in middle age, when people have already had children, there is little selective pressure on it to die out naturally. Indeed, in several studies, those with the mutations appear to breed more prolifically than their unaffected siblings.[3]

Huntington's was the first completely dominant human genetic disease to come to light. That means it is not like alkaptonuria in which you must have two copies of the mutant gene, one from each parent, to suffer the symptoms. Just one copy of the mutation will do. The disease seems to be worse if inherited from the father and the mutation tends to grow more severe, by the lengthening of the repeat, in the children of progressively older fathers.

In the late 1970s, a determined woman set out to find the Huntington gene. Following Woody Guthrie's terrible death from the disease, his widow started the Committee to Combat Huntington's Chorea; she was joined by a doctor named Milton Wexler whose wife and three brothers-in-law were suffering from the disease. Wexler's daughter, Nancy, knew she stood a fifty per cent chance of having the mutation herself and she became obsessed with finding the gene. She was told not to bother. The gene would prove impossible to find. It would be like looking for a needle in a haystack the size of America. She should wait a few years until the techniques were better and there was a realistic chance. 'But', she wrote, 'if you have Huntington's disease, you do not have time to wait.' Acting on the report of a Venezuelan doctor, Americo Negrette, in 1979 she flew to Venezuela to visit three rural villages called San Luis, Barranquitas and Laguneta on the shores of Lake Maracaibo. Actually a huge, almost landlocked gulf of the sea, Lake Maracaibo lies in the far west of Venezuela, beyond the Cordillera de Merida.

The area contained a vast, extended family with a high incidence of Huntington's disease. The story they told each other was that the affliction came from an eighteenth-century sailor, and Wexler was able to trace the family tree of the disease back to the early nineteenth century and a woman called, appropriately, Maria Concepcion. She lived in the Pueblos de Agua, villages of houses built

on stilts over the water. A fecund ancestor, she had 11,000 descendants in eight generations, 9,000 of whom were still alive in 1981. No less than 371 of them had Huntington's disease when Wexler first visited and 3,600 carried a risk of at least a quarter that they would develop the disease, because at least one grandparent had the symptoms.

Wexler's courage was extraordinary, given that she too might have the mutation. 'It is crushing to look at these exuberant children', she wrote,[4] 'full of hope and expectation, despite poverty, despite illiteracy, despite dangerous and exhausting work for the boys fishing in small boats in the turbulent lake, or for even the tiny girls tending house and caring for ill parents, despite a brutalising disease robbing them of parents, grandparents, aunts, uncles, and cousins – they are joyous and wild with life, until the disease attacks.'

Wexler started searching the haystack. First she collected blood from over 500 people: 'hot, noisy days of drawing blood'. Then she sent it to Jim Gusella's laboratory in Boston. He began to test genetic markers in search of the gene: randomly chosen chunks of DNA, that might or might not turn out to be reliably different in the affected and unaffected people. Fortune smiled on him and by mid-1983 he had not only isolated a marker close to the gene affected, but pinned it down to the tip of the short arm of chromosome 4. He knew which three-millionth of the genome it was in. Home and dry? Not so fast. The gene lay in a region of the text one million 'letters' long. The haystack was smaller, but still vast. Eight years later the gene was still mysterious: 'The task has been arduous in the extreme', wrote Wexler,[4] sounding like a Victorian explorer, 'in this inhospitable terrain at the top of chromosome 4. It has been like crawling up Everest over the past eight years.'

The persistence paid off. In 1993, the gene was found at last, its text was read and the mutation that led to the disease identified. The gene is the recipe for a protein called huntingtin: the protein was discovered after the gene – hence its name. The repetition of the 'word' CAG in the middle of the gene results in a long stretch of glutamines in the middle of the protein (CAG means glutamine in 'genetish'). And, in the case of Huntington's disease, the more glutamines there are at this point, the earlier in life the disease begins.[5]

It seems a desperately inadequate explanation of the disease. If the huntingtin gene is damaged, then why does it work all right for the first thirty years of life? Apparently, the mutant form of huntingtin very gradually accumulates in aggregate chunks. Like Alzheimer's disease and BSE, it is this accumulation of a sticky lump of protein within the cell that causes the death of the cell, perhaps because it induces the cell to commit suicide. In Huntington's disease this happens mostly within the brain's dedicated movement-control room, the cerebellum, with the result that movement becomes progressively less easy or controlled.[6]

The most unexpected feature of the stuttering repetition of the word CAG is that it is not confined to Huntington's disease. There are five other neurological diseases caused by so-called 'unstable CAG repeats' in entirely different genes. Cerebellar ataxia is one. There is even a bizarre report that a long CAG repeat deliberately inserted into a random gene in a mouse caused a late-onset, neurological disease rather like Huntington's disease. CAG repeats may therefore cause neurological disease whatever the gene in which they appear. Moreover, there are other diseases of nerve degeneration caused by other stuttering repeats of 'words' and in every case the repeated 'word' begins with C and ends in G. Six different CAG diseases are known. CCG or CGG repeated more than 200 times near the beginning of a gene on the X chromosome causes 'fragile X', a variable but unusually common form of mental retardation (fewer than sixty repeats is normal; up to a thousand is possible). CTG repeated from fifty to one thousand times in a gene on chromosome 19 causes myotonic dystrophy. More than a dozen human diseases are caused by expanded three-letter word repeats – the so-called polyglutamine diseases. In all cases the elongated protein has a tendency to accumulate in indigestible lumps that cause their cells to die. The different symptoms are caused by the fact that different genes are switched on in different parts of the body.[7]

What is so special about the 'word' C*G, apart from the fact that

it means glutamine? A clue comes from a phenomenon known as anticipation. It has been known for some time that those with a severe form of Huntington's disease or fragile X are likely to have children in whom the disease is worse or begins earlier than it did in themselves. Anticipation means that the longer the repetition, the longer it is likely to grow when copied for the next generation. We know that these repeats form little loopings of DNA called hairpins. The DNA likes to stick to itself, forming a structure like a hairpin, with the Cs and Gs of the C*G 'words' sticking together across the pin. When the hairpins unfold, the copying mechanism can slip and more copies of the word insert themselves.[8]

A simple analogy might be helpful. If I repeat a word six times in this sentence – cag, cag, cag, cag, cag, cag – you will count it fairly easily. But if I repeat it thirty-six times – cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag, cag – I am willing to bet you lose count. So it is with the DNA. The more repeats there are, the more likely the copying mechanism is to insert an extra one. Its finger slips and loses its place in the text. An alternative (or possibly additional) explanation is that the checking system, called mismatch repair, is good at catching small changes, but not big ones in C*G repeats.[9]

This may explain why the disease develops late in life. Laura Mangiarini at Guy's Hospital in London created transgenic mice, equipped with copies of part of the Huntington's gene that contained more than one hundred repeats. As the mice grew older, so the length of the gene increased in all their tissues save one. Up to ten extra CAG 'words' were added to it. The one exception was the cerebellum, the hindbrain responsible for controlling movement. The cells of the cerebellum do not need to change during life once the mice have learnt to walk, so they never divide. It is when cells and genes divide that copying mistakes are made. In human beings, the number of repeats in the cerebellum *falls* during life, though it increases in other tissues. In the cells from which sperm are made, the CAG repeats grow, which explains why there is a relationship between the onset of Huntington's disease and the age of the father: older fathers have sons who get the disease more severely and at a younger age. (Incidentally, it is now known that the mutation rate, throughout the genome, is about five times as high in men as it is in women, because of the repeated replication needed to supply fresh sperm cells throughout life.)[10]

Some families seem to be more prone to the spontaneous appearance of the Huntington's mutation than others. The reason seems to be not only that they have a repeat number just below the threshold (say between twenty-nine and thirty-five), but that it jumps above the threshold about twice as easily as it does in other people with similar repeat numbers. The reason for that is again a simple matter of letters. Compare two people: one has thirty-five CAGs followed by a bunch of CCAs and CCGs. If the reader slips and adds an extra CAG, the repeat number grows by one. The other person has thirty-five CAGs, followed by a CAA then two more CAGs. If the reader slips and misreads the CAA as a CAG, the effect is to add not one but three to the repeat number, because of the two CAGs already waiting.[11]

Though I seem to be getting carried away, and deluging you with details about CAGs in the huntingtin gene, consider: almost none of this was known five years ago. The gene had not been found, the CAG repeat had not been identified, the huntingtin protein was unknown, the link with other neurodegenerative diseases was not even guessed at, the mutation rates and causes were mysterious, the paternal age effect was unexplained. From 1872 to 1993 virtually nothing was known about Huntington's disease except that it was genetic. This mushroom of knowledge has grown up almost overnight since then, a mushroom vast enough to require days in a library merely to catch up. The number of scientists who have published papers on the Huntington's gene since 1993 is close to 100. All about one gene. One of 60,000–80,000 genes in the human genome. If you still need convincing of the immensity of the Pandora's box that James Watson and Francis Crick opened that day in 1953, the Huntington's story will surely persuade you. Compared

with the knowledge to be gleaned from the genome, the whole of the rest of biology is but a thimbleful.

And yet not a single case of Huntington's disease has been cured. The knowledge that I celebrate has not even suggested a remedy for the affliction. If anything, in the heartless simplicity of the CAG repeats, it has made the picture look even bleaker for those seeking a cure. There are 100 billion cells in the brain. How can we go in and shorten the CAG repeats in the huntingtin genes of each and every one?

Nancy Wexler relates a story about a woman in the Lake Maracaibo study. She came to Wexler's hut to be tested for neurological signs of the disease. She seemed fine and well but Wexler knew that small hints of Huntington's can be detected by certain tests long before the patient herself sees signs. Sure enough this woman showed such signs. But unlike most people, when the doctors had finished their examination, she asked them what their conclusion was. Did she have the disease? The doctor replied with a question: What do you think? She thought she was all right. The doctors avoided saying what they thought, mentioning the need to get to know people better before they gave diagnoses. As soon as the woman left the room, her friend came rushing in, almost hysterical. What did you tell her? The doctors recounted what they had said. 'Thank God', replied the friend and explained: the woman had said to the friend that she would ask for the diagnosis and if it turned out that she had Huntington's disease, she would immediately go and commit suicide.

There are several things about that story that are disturbing. The first is the falsely happy ending. The woman does have the mutation. She faces a death sentence, whether by her hand or much more slowly. She cannot escape her fate, however nicely she is treated by the experts. And surely the knowledge about her condition is hers to do with as she wishes. If she wishes to act on it and kill herself, who are the doctors to withhold the information? Yet they did the 'right thing', too. Nothing is more sensitive than the results of a test for a fatal disease; telling people the result starkly and coldly

may well not be the best thing to do – for them. Testing without counselling is a recipe for misery. But above all the tale drives home the uselessness of diagnosing without curing. The woman thought she was all right. Suppose she had five more years of happy ignorance ahead of her; there is no point in telling her that after that she faces lurching madness.

A person who has watched her mother die from Huntington's disease knows she has a fifty per cent chance of contracting it. But that is not right, is it? No individual can have fifty per cent of this disease. She either has a one hundred per cent chance or zero chance, and the probability of each is equal. So all that a genetic test does is unpackage the risk and tell her whether her ostensible fifty per cent is actually one hundred per cent or is actually zero.

Nancy Wexler fears that science is now in the position of Tiresias, the blind seer of Thebes. By accident Tiresias saw Athena bathing and she struck him blind. Afterwards she repented and, unable to restore his sight, gave him the power of soothsaying. But seeing the future was a terrible fate, since he could see it but not change it. 'It is but sorrow', said Tiresias to Oedipus, 'to be wise when wisdom profits not.' Or as Wexler puts it, 'Do you want to know when you are going to die, especially if you have no power to change the outcome?' Many of those at risk from Huntington's disease, who since 1986 can have themselves tested for the mutation, choose ignorance. Only about twenty per cent of them choose to take the test. Curiously, but perhaps understandably, men are three times as likely to choose ignorance as women. Men are more concerned with themselves rather than their progeny.[12]

Even if those at risk choose to know, the ethics are byzantine. If one member of a family takes the test, he or she is in effect testing the whole family. Many parents take the test reluctantly but for the sake of their children. And misconceptions abound, even in textbooks and medical leaflets. Half your children may suffer, says one, addressing parents with the mutation. Not so: each child has a fifty per cent chance, which is very different. How the result of the test is presented is also immensely sensitive. Psychologists have

found that people feel better about being told they have a three-quarter chance of an unaffected baby than if they are told they have a one-quarter chance of an affected one. Yet they are the same thing.

Huntington's disease is at the far end of a spectrum of genetics. It is pure fatalism, undiluted by environmental variability. Good living, good medicine, healthy food, loving families or great riches can do nothing about. Your fate is in your genes. Like a pure Augustinian, you go to heaven by God's grace, not by good works. It reminds us that the genome, great book that it is, may give us the bleakest kind of self-knowledge: the knowledge of our destiny, not the kind of knowledge that you can do something about, but the curse of Tiresias.

Yet Nancy Wexler's obsession with finding the gene was driven by her desire to mend it or cure it when she did find it. And she is undoubtedly closer to that goal now than ten years ago. 'I am an optimist', she writes,[4] 'Even though I feel this hiatus in which we will be able only to predict and not to prevent will be exceedingly difficult . . . I believe the knowledge will be worth the risks.'

What of Nancy Wexler herself? Several times in the late 1980s, she and her elder sister Alice sat down with their father Milton to discuss whether either of the women should take the test. The debates were tense, angry and inconclusive. Milton was against taking the test, stressing its uncertainties and the danger of a false diagnosis. Nancy had been determined that she wanted the test, but her determination gradually evaporated in the face of a real possibility. Alice chronicled the discussions in a diary that later became a soul-searching book called *Mapping fate*. The result was that neither woman took the test. Nancy is now the same age as her mother was when she was diagnosed.[13]

CHROMOSOME 5

Environment

Errors, like straws, upon the surface flow;
He who would search for pearls must dive below.
John Dryden, All for Love

It is time for a cold shower. Reader, the author of this book has been misleading you. He has repeatedly used the word 'simple' and burbled on about the surprising simplicity at the heart of genetics. A gene is just a sentence of prose written in a very simple language, he says, preening himself at the metaphor. Such a simple gene on chromosome 3 is the cause, when broken, of alkaptonuria. Another gene on chromosome 4 is the cause, when elongated, of Huntington's chorea. You either have mutations, in which case you get these genetic diseases, or you don't. No need for waffle, statistics or fudge. It is a digital world, this genetics stuff, all particulate inheritance. Your peas are either wrinkled or they are smooth.

You have been misled. The world is not like that. It is a world of greys, of nuances, of qualifiers, of 'it depends'. Mendelian genetics is no more relevant to understanding heredity in the real world than Euclidean geometry is to understanding the shape of an oak tree.