The Selfish Gene

RICHARD DAWKINS

NEW EDITION

Oxford New York
OXFORD UNIVERSITY PRESS
1989

3

Immortal coils

We are survival machines, but 'we' does not mean just people. It embraces all animals, plants, bacteria, and viruses. The total number of survival machines on earth is very difficult to count and even the total number of species is unknown. Taking just insects alone, the number of living species has been estimated at around three million, and the number of individual insects may be a million million million.

Different sorts of survival machine appear very varied on the outside and in their internal organs. An octopus is nothing like a mouse, and both are quite different from an oak tree. Yet in their fundamental chemistry they are rather uniform, and, in particular, the replicators that they bear, the genes, are basically the same kind of molecule in all of us—from bacteria to elephants. We are all survival machines for the same kind of replicator—molecules called DNA—but there are many different ways of making a living in the world, and the replicators have built a vast range of machines to exploit them. A monkey is a machine that preserves genes up trees, a fish is a machine that preserves genes in the water; there is even a small worm that preserves genes in German beer mats. DNA works in mysterious ways.

For simplicity I have given the impression that modern genes, made of DNA, are much the same as the first replicators in the primeval soup. It does not matter for the argument, but this may not really be true. The original replicators may have been a related kind of molecule to DNA, or they may have been totally different. In the latter case we might say that their survival machines must have been seized at a later stage by DNA. If so, the original replicators were utterly destroyed, for no trace of them remains in modern survival machines. Along these lines, A. G. Cairns-Smith has made the intriguing suggestion that our ancestors, the first replicators, may have been not organic molecules at all, but inorganic crystals—

minerals, little bits of clay. Usurper or not, DNA is in undisputed charge today, unless, as I tentatively suggest in Chapter 11, a new seizure of power is now just beginning.

A DNA molecule is a long chain of building blocks, small molecules called nucleotides. Just as protein molecules are chains of amino acids, so DNA molecules are chains of nucleotides. A DNA molecule is too small to be seen, but its exact shape has been ingeniously worked out by indirect means. It consists of a pair of nucleotide chains twisted together in an elegant spiral; the 'double helix'; the 'immortal coil'. The nucleotide building blocks come in only four different kinds, whose names may be shortened to *A*, *T*, *C*, and *G*. These are the same in all animals and plants. What differs is the order in which they are strung together. A *G* building block from a man is identical in every particular to a *G* building block from a snail. But the *sequence* of building blocks in a man is not only different from that in a snail. It is also different—though less so—from the sequence in every other man (except in the special case of identical twins).

Our DNA lives inside our bodies. It is not concentrated in a particular part of the body, but is distributed among the cells. There are about a thousand million million cells making up an average human body, and, with some exceptions which we can ignore, every one of those cells contains a complete copy of that body's DNA. This DNA can be regarded as a set of instructions for how to make a body, written in the *A*, *T*, *C*, *G* alphabet of the nucleotides. It is as though, in every room of a gigantic building, there was a book-case containing the architect's plans for the entire building. The 'book-case' in a cell is called the nucleus. The architect's plans run to 46 volumes in man—the number is different in other species. The 'volumes' are called chromosomes. They are visible under a microscope as long threads, and the genes are strung out along them in order. It is not easy, indeed it may not even be meaningful, to decide where one gene ends and the next one begins. Fortunately, as this chapter will show, this does not matter for our purposes.

I shall make use of the metaphor of the architect's plans, freely mixing the language of the metaphor with the language of the real thing. 'Volume' will be used interchangeably with chromosome. 'Page' will provisionally be used interchangeably with gene, although the division between genes is less clear-cut than the division between the pages of a book. This metaphor will take us quite a long way.

When it finally breaks down I shall introduce other metaphors. Incidentally, there is of course no 'architect'. The DNA instructions have been assembled by natural selection.

DNA molecules do two important things. Firstly they replicate, that is to say they make copies of themselves. This has gone on non-stop ever since the beginning of life, and the DNA molecules are now very good at it indeed. As an adult, you consist of a thousand million million cells, but when you were first conceived you were just a single cell, endowed with one master copy of the architect's plans. This cell divided into two, and each of the two cells received its own copy of the plans. Successive divisions took the number of cells up to 4, 8, 16, 32, and so on into the billions. At every division the DNA plans were faithfully copied, with scarcely any mistakes.

It is one thing to speak of the duplication of DNA. But if the DNA is really a set of plans for building a body, how are the plans put into practice? How are they translated into the fabric of the body? This brings me to the second important thing DNA does. It indirectly supervises the manufacture of a different kind of molecule—protein. The haemoglobin which was mentioned in the last chapter is just one example of the enormous range of protein molecules. The coded message of the DNA, written in the four-letter nucleotide alphabet, is translated in a simple mechanical way into another alphabet. This is the alphabet of amino acids which spells out protein molecules.

Making proteins may seem a far cry from making a body, but it is the first small step in that direction. Proteins not only constitute much of the physical fabric of the body; they also exert sensitive control over all the chemical processes inside the cell, selectively turning them on and off at precise times and in precise places. Exactly how this eventually leads to the development of a baby is a story which it will take decades, perhaps centuries, for embryologists to work out. But it is a fact that it does. Genes do indirectly control the manufacture of bodies, and the influence is strictly one way: acquired characteristics are not inherited. No matter how much knowledge and wisdom you acquire during your life, not one jot will be passed on to your children by genetic means. Each new generation starts from scratch. A body is the genes' way of preserving the genes unaltered.

The evolutionary importance of the fact that genes control embryonic development is this: it means that genes are at least partly responsible for their own survival in the future, because their survival

depends on the efficiency of the bodies in which they live and which they helped to build. Once upon a time, natural selection consisted of the differential survival of replicators floating free in the primeval soup. Now, natural selction favours replicators that are good at building survival machines, genes that are skilled in the art of controlling embryonic development. In this, the replicators are no more conscious or purposeful than they ever were. The same old processes of automatic selection between rival molecules by reason of their longevity, fecundity, and copying-fidelity, still go on as blindly and as inevitably as they did in the far-off days. Genes have no foresight. They do not plan ahead. Genes just *are*, some genes more so than others, and that is all there is to it. But the qualities that determine a gene's longevity and fecundity are not so simple as they were. Not by a long way.

In recent years—the last six hundred million or so—the replicators have achieved notable triumphs of survival-machine technology such as the muscle, the heart, and the eye (evolved several times independently). Before that, they radically altered fundamental features of their way of life as replicators, which must be understood if we are to proceed with the argument.

The first thing to grasp about a modern replicator is that it is highly gregarious. A survival machine is a vehicle containing not just one gene but many thousands. The manufacture of a body is a cooperative venture of such intricacy that it is almost impossible to disentangle the contribution of one gene from that of another.* A given gene will have many different effects on quite different parts of the body. A given part of the body will be influenced by many genes, and the effect of any one gene depends on interaction with many others. Some genes act as master genes controlling the operation of a cluster of other genes. In terms of the analogy, any given page of the plans makes reference to many different parts of the building; and each page makes sense only in terms of cross-references to numerous other pages.

This intricate inter-dependence of genes may make you wonder why we use the word 'gene' at all. Why not use a collective noun like 'gene complex'? The answer is that for many purposes that is indeed quite a good idea. But if we look at things in another way, it does make sense too to think of the gene complex as being divided up into discrete replicators or genes. This arises because of the phenomenon of sex. Sexual reproduction has the effect of mixing

and shuffling genes. This means that any one individual body is just a temporary vehicle for a short-lived combination of genes. The *combination* of genes that is any one individual may be short-lived, but the genes themselves are potentially very long-lived. Their paths constantly cross and recross down the generations. One gene may be regarded as a unit that survives through a large number of successive individual bodies. This is the central argument that will be developed in this chapter. It is an argument that some of my most respected colleagues obstinately refuse to agree with, so you must forgive me if I seem to labour it! First I must briefly explain the facts of sex.

I said that the plans for building a human body are spelt out in 46 volumes. In fact this was an over-simplification. The truth is rather bizarre. The 46 chromosomes consist of 23 *pairs* of chromosomes. We might say that, filed away in the nucleus of every cell, are two alternative sets of 23 volumes of plans. Call them Volume 1a and 1b, Volume 2a and Volume 2b etc., down to Volume 23a and Volume 23b. Of course the identifying numbers I use for volumes and, later, pages, are purely arbitrary.

We receive each chromosome intact from one of our two parents, in whose testis or ovary it was assembled. Volumes 1a, 2a, 3a, . . . came, say, from the father. Volumes 1b, 2b, 3b, . . . came from the mother. It is very difficult in practice, but in theory you could look with a microscope at the 46 chromosomes in any one of your cells, and pick out the 23 that came from your father and the 23 that came from your mother.

The paired chromosomes do not spend all their lives physically in contact with each other, or even near each other. In what sense then are they 'paired'? In the sense that each volume coming originally from the father can be regarded, page for page, as a direct alternative to one particular volume coming originally from the mother. For instance, Page 6 of Volume 13a and Page 6 of Volume 13b might both be 'about' eye colour; perhaps one says 'blue' while the other says 'brown'.

Sometimes the two alternative pages are identical, but in other cases, as in our example of eye colour, they differ. If they make contradictory 'recommendations', what does the body do? The answer varies. Sometimes one reading prevails over the other. In the eye colour example just given, the person would actually have brown eyes: the instructions for making blue eyes would be ignored in the

building of the body, though this does not stop them being passed on to future generations. A gene that is ignored in this way is called *recessive*. The opposite of a recessive gene is a *dominant* gene. The gene for brown eyes is dominant to the gene for blue eyes. A person has blue eyes only if both copies of the relevant page are unanimous in recommending blue eyes. More usually when two alternative genes are not identical, the result is some kind of compromise—the body is built to an intermediate design or something completely different.

When two genes, like the brown eye and the blue eye gene, are rivals for the same slot on a chromosome, they are called *alleles* of each other. For our purposes, the word allele is synonymous with rival. Imagine the volumes of architects' plans as being loose-leaf binders, whose pages can be detached and interchanged. Every Volume 13 must have a Page 6, but there are several possible Page 6s which could go in the binder between Page 5 and Page 7. One version says 'blue eyes', another possible version says 'brown eyes'; there may be yet other versions in the population at large which spell out other colours like green. Perhaps there are half a dozen alternative alleles sitting in the Page 6 position on the 13th chromosomes scattered around the population as a whole. Any given person only has two Volume 13 chromosomes. Therefore he can have a maximum of two alleles in the Page 6 slot. He may, like a blue-eyed person, have two copies of the same allele, or he may have any two alleles chosen from the half dozen alternatives available in the population at large.

You cannot, of course, literally go and choose your genes from a pool of genes available to the whole population. At any given time all the genes are tied up inside individual survival machines. Our genes are doled out to us at conception, and there is nothing we can do about this. Nevertheless, there is a sense in which, in the long term, the genes of the population in general can be regarded as a *gene pool*. This phrase is in fact a technical term used by geneticists. The gene pool is a worthwhile abstraction because sex mixes genes up, albeit in a carefully organized way. In particular, something like the detaching and interchanging of pages and wads of pages from loose-leaf binders really does go on, as we shall presently see.

I have described the normal division of a cell into two new cells, each one receiving a complete copy of all 46 chromosomes. This normal cell division is called *mitosis*. But there is another kind of cell

division called *meiosis*. This occurs only in the production of the sex cells; the sperms or eggs. Sperms and eggs are unique among our cells in that, instead of containing 46 chromosomes, they contain only 23. This is, of course, exactly half of 46—convenient when they fuse in sexual fertilization to make a new individual! Meiosis is a special kind of cell division, taking place only in testicles and ovaries, in which a cell with the full double set of 46 chromosomes divides to form sex cells with the single set of 23 (all the time using the human numbers for illustration).

A sperm, with its 23 chromosomes, is made by the meiotic division of one of the ordinary 46-chromosome cells in the testicle. Which 23 are put into any given sperm cell? It is clearly important that a sperm should not get just any old 23 chromosomes: it mustn't end up with two copies of Volume 13 and none of Volume 17. It would theoretically be possible for an individual to endow one of his sperms with chromosomes which came, say, entirely from his mother; that is Volume 1b, 2b, 3b, . . ., 23b. In this unlikely event, a child conceived by the sperm would inherit half her genes from her paternal grandmother, and none from her paternal grandfather. But in fact this kind of gross, whole-chromosome distribution does not happen. The truth is rather more complex. Remember that the volumes (chromosomes) are to be thought of as loose-leaf binders. What happens is that, during the manufacture of the sperm, single pages, or rather multi-page chunks, are detached and swapped with the corresponding chunks from the alternative volume. So, one particular sperm cell might make up its Volume 1 by taking the first 65 pages from Volume 1a, and pages 66 to the end from Volume 1b. This sperm cell's other 22 volumes would be made up in a similar way. Therefore every sperm cell made by an individual is unique, even though all his sperms assembled their 23 chromosomes from bits of the same set of 46 chromosomes. Eggs are made in a similar way in ovaries, and they too are all unique.

The real-life mechanics of this mixing are fairly well understood. During the manufacture of a sperm (or egg), bits of each paternal chromosome physically detach themselves and change places with exactly corresponding bits of maternal chromosome. (Remember that we are talking about chromosomes that came originally from the parents of the individual making the sperm, i.e., from the paternal grandparents of the child who is eventually conceived by the sperm). The process of swapping bits of chromosome is called *crossing over*. It

is very important for the whole argument of this book. It means that if you got out your microscope and looked at the chromosomes in one of your own sperms (or eggs if you are female) it would be a waste of time trying to identify chromosomes that originally came from your father and chromosomes that originally came from your mother. (This is in marked contrast to the case of ordinary body cells (see page 25).) Any one chromosome in a sperm would be a patchwork, a mosaic of maternal genes and paternal genes.

The metaphor of the page for the gene starts to break down here. In a loose-leaf binder a whole page may be inserted, removed or exchanged, but not a fraction of a page. But the gene complex is just a long string of nucleotide letters, not divided into discrete pages in an obvious way at all. To be sure, there are special symbols for END OF PROTEIN CHAIN MESSAGE and START OF PROTEIN CHAIN MESSAGE written in the same four-letter alphabet as the protein messages themselves. In between these two punctuation marks are the coded instructions for making one protein. If we wish, we can define a single gene as a sequence of nucleotide letters lying between a START and an END symbol, and coding for one protein chain. The word *cistron* has been used for a unit defined in this way, and some people use the word gene interchangeably with cistron. But crossing-over does not respect boundaries between cistrons. Splits may occur within cistrons as well as between them. It is as though the architect's plans were written out, not on discrete pages, but on 46 rolls of ticker tape. Cistrons are not of fixed length. The only way to tell where one cistron ends and the next begins would be to read the symbols on the tape, looking for END OF MESSAGE and START OF MESSAGE symbols. Crossing-over is represented by taking matching paternal and maternal tapes, and cutting and exchanging matching portions, regardless of what is written on them.

In the title of this book the word gene means not a single cistron but something more subtle. My definition will not be to everyone's taste, but there is no universally agreed definition of a gene. Even if there were, there is nothing sacred about definitions. We can define a word how we like for our own purposes, provided we do so clearly and unambiguously. The definition I want to use comes from G. C. Williams.* A gene is defined as any portion of chromosomal material that potentially lasts for enough generations to serve as a unit of natural selection. In the words of the previous chapter, a gene is a replicator with high copying-fidelity. Copying-fidelity is another way

of saying longevity-in-the-form-of-copies and I shall abbreviate this simply to longevity. The definition will take some justifying.

On any definition, a gene has to be a portion of a chromosome. The question is, how big a portion—how much of the ticker tape? Imagine any sequence of adjacent code-letters on the tape. Call the sequence a *genetic unit*. It might be a sequence of only ten letters within one cistron; it might be a sequence of eight cistrons; it might start and end in mid-cistron. It will overlap with other genetic units. It will include smaller units, and it will form part of larger units. No matter how long or short it is, for the purposes of the present argument, this is what we are calling a genetic unit. It is just a length of chromosome, not physically differentiated from the rest of the chromosome in any way.

Now comes the important point. The shorter a genetic unit is, the longer—in generations—it is likely to live. In particular, the less likely it is to be split by any one crossing-over. Suppose a whole chromosome is, on average, likely to undergo one cross-over every time a sperm or egg is made by meiotic division, and this cross-over can happen anywhere along its length. If we consider a very large genetic unit, say half the length of the chromosome, there is a 50 per cent chance that the unit will be split at each meiosis. If the genetic unit we are considering is only 1 per cent of the length of the chromosome, we can assume that it has only a 1 per cent chance of being split in any one meiotic division. This means that the unit can expect to survive for a large number of generations in the individual's descendants. A single cistron is likely to be much less than 1 per cent of the length of a chromosome. Even a group of several neighbouring cistrons can expect to live many generations before being broken up by crossing over.

The average life-expectancy of a genetic unit can conveniently be expressed in generations, which can in turn be translated into years. If we take a whole chromosome as our presumptive genetic unit, its life story lasts for only one generation. Suppose it is your chromosome number 8a, inherited from your father. It was created inside one of your father's testicles, shortly before you were conceived. It had never existed before in the whole history of the world. It was created by the meiotic shuffling process, forged by the coming together of pieces of chromosome from your paternal grandmother and your paternal grandfather. It was placed inside one particular sperm, and it was unique. The sperm was one of several millions, a

vast armada of tiny vessels, and together they sailed into your mother. This particular sperm (unless you are a non-identical twin) was the only one of the flotilla which found harbour in one of your mother's eggs—that is why you exist. The genetic unit we are considering, your chromosome number 8a, set about replicating itself along with all the rest of your genetic material. Now it exists, in duplicate form, all over your body. But when you in your turn come to have children, this chromosome will be destroyed when you manufacture eggs (or sperms). Bits of it will be interchanged with bits of your maternal chromosome number 8b. In any one sex cell, a new chromosome number 8 will be created, perhaps 'better' than the old one, perhaps 'worse', but, barring a rather improbable coincidence, definitely different, definitely unique. The life-span of a chromosome is one generation.

What about the life-span of a smaller genetic unit, say 1/100 of the length of your chromosome 8a? This unit too came from your father, but it very probably was not originally assembled in him. Following the earlier reasoning, there is a 99 per cent chance that he received it intact from one of his two parents. Suppose it was from his mother, your paternal grandmother. Again, there is a 99 per cent chance that she inherited it intact from one of her parents. Eventually, if we trace the ancestry of a small genetic unit back far enough, we will come to its original creator. At some stage it must have been created for the first time inside a testicle or an ovary of one of your ancestors.

Let me repeat the rather special sense in which I am using the word 'create'. The smaller sub-units which make up the genetic unit we are considering may well have existed long before. Our genetic unit was created at a particular moment only in the sense that the particular *arrangement* of sub-units by which it is defined did not exist before that moment. The moment of creation may have occurred quite recently, say in one of your grandparents. But if we consider a very small genetic unit, it may have been first assembled in a much more distant ancestor, perhaps an ape-like pre-human ancestor. Moreover, a small genetic unit inside you may go on just as far into the future, passing intact through a long line of your descendants.

Remember too that an individual's descendants constitute not a single line but a branching line. Whichever of your ancestors it was who 'created' a particular short length of your chromosome 8a, he or she very likely has many other descendants besides you. One of your genetic units may also be present in your second cousin. It may be

present in me, and in the Prime Minister, and in your dog, for we all share ancestors if we go back far enough. Also the same small unit might be assembled several times independently by chance: if the unit is small, the coincidence is not too improbable. But even a close relative is unlikely to share a whole chromosome with you. The smaller a genetic unit is, the more likely it is that another individual shares it—the more likely it is to be represented many times over in the world, in the form of copies.

The chance coming together, through crossing-over, of previously existing sub-units is the usual way for a new genetic unit to be formed. Another way—of great evolutionary importance even though it is rare—is called *point mutation*. A point mutation is an error corresponding to a single misprinted letter in a book. It is rare, but clearly the longer a genetic unit is, the more likely it is to be altered by a mutation somewhere along its length.

Another rare kind of mistake or mutation which has important long-term consequences is called *inversion*. A piece of chromosome detaches itself at both ends, turns head over heels, and reattaches itself in the inverted position. In terms of the earlier analogy, this would necessitate some renumbering of pages. Sometimes portions of chromosomes do not simply invert, but become reattached in a completely different part of the chromosome, or even join up with a different chromosome altogether. This corresponds to the transfer of a wad of pages from one volume to another. The importance of this kind of mistake is that, though usually disastrous, it can occasionally lead to the close *linkage* of pieces of genetic material which happen to work well together. Perhaps two cistrons which have a beneficial effect only when they are both present—they complement or reinforce each other in some way—will be brought close to each other by means of inversion. Then natural selection may tend to favour the new 'genetic unit' so formed, and it will spread through the future population. It is possible that gene complexes have, over the years, been extensively rearranged or 'edited' in this kind of way.

One of the neatest examples of this concerns the phenomenon known as *mimicry*. Some butterflies taste nasty. They are usually brightly and distinctively coloured, and birds learn to avoid them by their 'warning' marks. Now other species of butterfly that do not taste nasty cash in. They *mimic* the nasty ones. They are born looking like them in colour and shape (but not taste). They frequently fool

human naturalists, and they also fool birds. A bird who has once tasted a genuinely nasty butterfly tends to avoid all butterflies that look the same. This includes the mimics, and so genes for mimicry are favoured by natural selection. That is how mimicry evolves.

There are many different species of 'nasty' butterfly and they do not all look alike. A mimic cannot resemble all of them: it has to commit itself to one particular nasty species. In general, any particular species of mimic is a specialist at mimicking one particular nasty species. But there are species of mimic that do something very strange. Some individuals of the species mimic one nasty species; other individuals mimic another. Any individual who was intermediate or who tried to mimic both would soon be eaten; but such intermediates are not born. Just as an individual is either definitely male or definitely female, so an individual butterfly mimics either one nasty species or the other. One butterfly may mimic species *A* while his brother mimics species *B*.

It looks as though a single gene determines whether an individual will mimic species *A* or species *B*. But how can a single gene determine all the multifarious aspects of mimicry—colour, shape, spot pattern, rhythm of flight? The answer is that one gene in the sense of a *cistron* probably cannot. But by the unconscious and automatic 'editing' achieved by inversions and other accidental rearrangements of genetic material, a large cluster of formerly separate genes has come together in a tight linkage group on a chromosome. The whole cluster behaves like a single gene—indeed, by our definition it now *is* a single gene—and it as an 'allele' which is really another cluster. One cluster contains the cistrons concerned with mimicking species *A*; the other those concerned with mimicking species *B*. Each cluster is so rarely split up by crossing-over that an intermediate butterfly is never seen in nature, but they do very occasionally turn up if large numbers of butterflies are bred in the laboratory.

I am using the word gene to mean a genetic unit that is small enough to last for a large number of generations and to be distributed around in the form of many copies. This is not a rigid all-or-nothing definition, but a kind of fading-out definition, like the definition of 'big' or 'old'. The more likely a length of chromosome is to be split by crossing-over, or altered by mutations of various kinds, the less it qualifies to be called a gene in the sense in which I am using the term. A cistron presumably qualifies, but so also do larger units. A dozen

cistrons may be so close to each other on a chromosome that for our purposes they constitute a single long-lived genetic unit. The butterfly mimicry cluster is a good example. As the cistrons leave one body and enter the next, as they board sperm or egg for the journey into the next generation, they are likely to find that the little vessel contains their close neighbours of the previous voyage, old shipmates with whom they sailed on the long odyssey from the bodies of distant ancestors. Neighbouring cistrons on the same chromosome form a tightly-knit troupe of travelling companions who seldom fail to get on board the same vessel when meiosis time comes around.

To be strict, this book should be called not *The Selfish Cistron* nor *The Selfish Chromosome*, but *The slightly selfish big bit of chromosome and the even more selfish little bit of chromosome*. To say the least this is not a catchy title so, defining a gene as a little bit of chromosome which potentially lasts for many generations, I call the book *The Selfish Gene*.

We have now arrived back at the point we left at the end of Chapter 1. There we saw that selfishness is to be expected in any entity that deserves the title of a basic unit of natural selection. We saw that some people regard the species as the unit of natural selection, others the population or group within the species, and yet others the individual. I said that I preferred to think of the gene as the fundamental unit of natural selection, and therefore the fundamental unit of self-interest. What I have now done is to *define* the gene in such a way that I cannot really help being right!

Natural selection in its most general form means the differential survival of entities. Some entities live and others die but, in order for this selective death to have any impact on the world, an additional condition must be met. Each entity must exist in the form of lots of copies, and at least some of the entities must be *potentially* capable of surviving—in the form of copies—for a significant period of evolutionary time. Small genetic units have these properties: individuals, groups, and species do not. It was the great achievement of Gregor Mendel to show that hereditary units can be treated in practice as indivisible and independent particles. Nowadays we know that this is a little too simple. Even a cistron is occasionally divisible and any two genes on the same chromosome are not wholly independent. What I have done is to define a gene as a unit which, to a high degree, *approaches* the ideal of indivisible particulateness. A gene is not

indivisible, but it is seldom divided. It is either definitely present or definitely absent in the body of any given individual. A gene travels intact from grandparent to grandchild, passing straight through the intermediate generation without being merged with other genes. If genes continually blended with each other, natural selection as we now understand it would be impossible. Incidentally, this was proved in Darwin's lifetime, and it caused Darwin great worry since in those days it was assumed that heredity was a blending process. Mendel's discovery had already been published, and it could have rescued Darwin, but alas he never knew about it: nobody seems to have read it until years after Darwin and Mendel had both died. Mendel perhaps did not realize the significance of his findings, otherwise he might have written to Darwin.

Another aspect of the particulateness of the gene is that it does not grow senile; it is no more likely to die when it is a million years old than when it is only a hundred. It leaps from body to body down the generations, manipulating body after body in its own way and for its own ends, abandoning a succession of mortal bodies before they sink in senility and death.

The genes are the immortals, or rather, they are defined as genetic entities that come close to deserving the title. We, the individual survival machines in the world, can expect to live a few more decades. But the genes in the world have an expectation of life that must be measured not in decades but in thousands and millions of years.

In sexually reproducing species, the individual is too large and too temporary a genetic unit to qualify as a significant unit of natural selection.* The group of individuals is an even larger unit. Genetically speaking, individuals and groups are like clouds in the sky or dust-storms in the desert. They are temporary aggregations or federations. They are not stable through evolutionary time. Populations may last a long while, but they are constantly blending with other populations and so losing their identity. They are also subject to evolutionary change from within. A population is not a discrete enough entity to be a unit of natural selection, not stable and unitary enough to be 'selected' in preference to another population.

An individual body seems discrete enough while it lasts, but alas, how long is that? Each individual is unique. You cannot get evolution by selecting between entities when there is only one copy of each entity! Sexual reproduction is not replication. Just as a population is

contaminated by other populations, so an individual's posterity is contaminated by that of his sexual partner. Your children are only half you, your grandchildren only a quarter you. In a few generations the most you can hope for is a large number of descendants, each of whom bears only a tiny portion of you—a few genes—even if a few do bear your surname as well.

Individuals are not stable things, they are fleeting. Chromosomes too are shuffled into oblivion, like hands of cards soon after they are dealt. But the cards themselves survive the shuffling. The cards are the genes. The genes are not destroyed by crossing-over, they merely change partners and march on. Of course they march on. That is their business. They are the replicators and we are their survival machines. When we have served our purpose we are cast aside. But genes are denizens of geological time: genes are forever.

Genes, like diamonds, are forever, but not quite in the same way as diamonds. It is an individual diamond crystal that lasts, as an unaltered pattern of atoms. DNA molecules don't have that kind of permanence. The life of any one physical DNA molecule is quite short—perhaps a matter of months, certainly not more than one lifetime. But a DNA molecule could theoretically live on in the form of *copies* of itself for a hundred million years. Moreover, just like the ancient replicators in the primeval soup, copies of a particular gene may be distributed all over the world. The difference is that the modern versions are all neatly packaged inside the bodies of survival machines.

What I am doing is emphasizing the potential near-immortality of a gene, in the form of copies, as its defining property. To define a gene as a single cistron is good for some purposes, but for the purposes of evolutionary theory it needs to be enlarged. The extent of the enlargement is determined by the purpose of the definition. We want to find the practical unit of natural selection. To do this we begin by identifying the properties that a successful unit of natural selection must have. In the terms of the last chapter, these are longevity, fecundity, and copying-fidelity. We then simply define a 'gene' as the largest entity which, at least potentially, has these properties. The gene is a long-lived replicator, existing in the form of many duplicate copies. It is not infinitely long-lived. Even a diamond is not literally everlasting, and even a cistron can be cut in two by crossing-over. The gene is defined as a piece of chromosome which

is sufficiently short for it to last, potentially, for *long enough* for it to function as a significant unit of natural selection.

Exactly how long is 'long enough'? There is no hard and fast answer. It will depend on how severe the natural selection 'pressure' is. That is, on how much more likely a 'bad' genetic unit is to die than its 'good' allele. This is a matter of quantitative detail which will vary from example to example. The largest practical unit of natural selection—the gene—will usually be found to lie somewhere on the scale between cistron and chromosome.

It is its potential immortality that makes a gene a good candidate as the basic unit of natural selection. But now the time has come to stress the word 'potential'. A gene *can* live for a million years, but many new genes do not even make it past their first generation. The few new ones that succeed do so partly because they are lucky, but mainly because they have what it takes, and that means they are good at making survival machines. They have an effect on the embryonic development of each successive body in which they find themselves, such that that body is a little bit more likely to live and reproduce than it would have been under the influence of the rival gene or allele. For example, a 'good' gene might ensure its survival by tending to endow the successive bodies in which it finds itself with long legs, which help those bodies to escape from predators. This is a particular example, not a universal one. Long legs, after all, are not always an asset. To a mole they would be a handicap. Rather than bog ourselves down in details, can we think of any *universal* qualities that we would expect to find in all good (i.e. long-lived) genes? Conversely, what are the properties that instantly mark a gene out as a 'bad', short-lived one? There might be several such universal properties, but there is one that is particularly relevant to this book: at the gene level, altruism must be bad and selfishness good. This follows inexorably from our definitions of altruism and selfishness. Genes are competing directly with their alleles for survival, since their alleles in the gene pool are rivals for their slot on the chromosomes of future generations. Any gene that behaves in such a way as to increase its own survival chances in the gene pool at the expense of its alleles will, by definition, tautologously, tend to survive. The gene is the basic unit of selfishness.

The main message of this chapter has now been stated. But I have glossed over some complications and hidden assumptions. The first complication has already been briefly mentioned. However

independent and free genes may be in their journey through the generations, they are very much *not* free and independent agents in their control of embryonic development. They collaborate and interact in inextricably complex ways, both with each other, and with their external environment. Expressions like 'gene for long legs' or 'gene for altruistic behaviour' are convenient figures of speech, but it is important to understand what they mean. There is no gene which single-handedly builds a leg, long or short. Building a leg is a multi-gene cooperative enterprise. Influences from the external environment too are indispensable: after all, legs are actually made of food! But there may well be a single gene which, *other things being equal*, tends to make legs longer than they would have been under the influence of the gene's allele.

As an analogy, think of the influence of a fertilizer, say nitrate, on the growth of wheat. Everybody knows that wheat plants grow bigger in the presence of nitrate than in its absence. But nobody would be so foolish as to claim that, on its own, nitrate can make a wheat plant. Seed, soil, sun, water, and various minerals are obviously all necessary as well. But if all these other factors are held constant, and even if they are allowed to vary within limits, addition of nitrate will make the wheat plants grow bigger. So it is with single genes in the development of an embryo. Embryonic development is controlled by an interlocking web of relationships so complex that we had best not contemplate it. No one factor, genetic or environmental, can be considered as the single 'cause' of any part of a baby. All parts of a baby have a near infinite number of antecedent causes. But a *difference* between one baby and another, for example a difference in length of leg, might easily be traced to one or a few simple antecedent differences, either in environment or in genes. It is *differences* that matter in the competitive struggle to survive; and it is genetically-controlled differences that matter in evolution.

As far as a gene is concerned, its alleles are its deadly rivals, but other genes are just a part of its environment, comparable to temperature, food, predators, or companions. The effect of the gene depends on its environment, and this includes other genes. Sometimes a gene has one effect in the presence of a particular other gene, and a completely different effect in the presence of another set of companion genes. The whole set of genes in a body constitutes a kind of genetic climate or background, modifying and influencing the effects of any particular gene.

But now we seem to have a paradox. If building a baby is such an intricate cooperative venture, and if every gene needs several thousands of fellow genes to complete its task, how can we reconcile this with my picture of indivisible genes, springing like immortal chamois from body to body down the ages: the free, untrammelled, and self-seeking agents of life? Was that all nonsense? Not at all. I may have got a bit carried away with the purple passages, but I was not talking nonsense, and there is no real paradox. We can explain this by means of another analogy.

One oarsman on his own cannot win the Oxford and Cambridge boat race. He needs eight colleagues. Each one is a specialist who always sits in a particular part of the boat—bow or stroke or cox etc. Rowing the boat is a cooperative venture, but some men are nevertheless better at it than others. Suppose a coach has to choose his ideal crew from a pool of candidates, some specializing in the bow position, others specializing as cox, and so on. Suppose that he makes his selection as follows. Every day he puts together three new trial crews, by random shuffling of the candidates for each position, and he makes the three crews race against each other. After some weeks of this it will start to emerge that the winning boat often tends to contain the same individual men. These are marked up as good oarsmen. Other individuals seem consistently to be found in slower crews, and these are eventually rejected. But even an outstandingly good oarsman might sometimes be a member of a slow crew, either because of the inferiority of the other members, or because of bad luck—say a strong adverse wind. It is only *on average* that the best men tend to be in the winning boat.

The oarsmen are genes. The rivals for each seat in the boat are alleles potentially capable of occupying the same slot along the length of a chromosome. Rowing fast corresponds to building a body which is successful at surviving. The wind is the external environment. The pool of alternative candidates is the gene pool. As far as the survival of any one body is concerned, all its genes are in the same boat. Many a good gene gets into bad company, and finds itself sharing a body with a lethal gene, which kills the body off in childhood. Then the good gene is destroyed along with the rest. But this is only one body, and replicas of the same good gene live on in other bodies which lack the lethal gene. Many copies of good genes are dragged under because they happen to share a body with bad genes, and many perish through other forms of ill luck, say when

their body is struck by lightning. But by definition luck, good and bad, strikes at random, and a gene that is *consistently* on the losing side is not unlucky; it is a bad gene.

One of the qualities of a good oarsman is teamwork, the ability to fit in and cooperate with the rest of a crew. This may be just as important as strong muscles. As we saw in the case of the butterflies, natural selection may unconsciously 'edit' a gene complex by means of inversions and other gross movements of bits of chromosome, thereby bringing genes that cooperate well together into closely linked groups. But there is also a sense in which genes which are in no way linked to each other physically can be selected for their mutual compatibility. A gene that cooperates well with most of the other genes that it is likely to meet in successive bodies, i.e. the genes in the whole of the rest of the gene pool, will tend to have an advantage.

For example, a number of attributes are desirable in an efficient carnivore's body, among them sharp cutting teeth, the right kind of intestine for digesting meat, and many other things. An efficient herbivore, on the other hand, needs flat grinding teeth, and a much longer intestine with a different kind of digestive chemistry. In a herbivore gene pool, any new gene that conferred on its possessors sharp meat-eating teeth would not be very successful. This is not because meat-eating is universally a bad idea, but because you cannot efficiently eat meat unless you also have the right sort of intestine, and all the other attributes of a meat-eating way of life. Genes for sharp, meat-eating teeth are not inherently bad genes. They are only bad genes in a gene-pool that is dominated by genes for herbivorous qualities.

This is a subtle, complicated idea. It is complicated because the 'environment' of a gene consists largely of other genes, each of which is itself being selected for its ability to cooperate with *its* environment of other genes. An analogy adequate to cope with this subtle point does exist, but it is not from everyday experience. It is the analogy with human 'game theory', which will be introduced in Chapter 5 in connection with aggressive contests between individual animals. I therefore postpone further discussion of this point until the end of that chapter, and return to the central message of this one. This is that the basic unit of natural selection is best regarded not as the species, nor as the population, nor even as the individual, but as some small unit of genetic material which it is convenient to label the gene.

The cornerstone of the argument, as given earlier, was the assumption that genes are potentially immortal, while bodies and all other higher units are temporary. This assumption rests upon two facts: the fact of sexual reproduction and crossing-over, and the fact of individual mortality. These facts are undeniably true. But this does not stop us asking why they are true. Why do we and most other survival machines practise sexual reproduction? Why do our chromosomes cross over? And why do we not live for ever?

The question of why we die of old age is a complex one, and the details are beyond the scope of this book. In addition to particular reasons, some more general ones have been proposed. For example, one theory is that senility represents an accumulation of deleterious copying errors and other kinds of gene damage which occur during the individual's lifetime. Another theory, due to Sir Peter Medawar, is a good example of evolutionary thinking in terms of gene selection.* Medawar first dismisses traditional arguments such as: 'Old individuals die as an act of altruism to the rest of the species, because if they stayed around when they were too decrepit to reproduce, they would clutter up the world to no good purpose.' As Medawar points out, this is a circular argument, assuming what it sets out to prove, namely that old animals are too decrepit to reproduce. It is also a naïve group-selection or species-selection kind of explanation, although that part of it could be rephrased more respectably. Medawar's own theory has a beautiful logic. We can build up to it as follows.

We have already asked what are the most general attributes of a 'good' gene, and we decided that 'selfishness' was one of them. But another general quality that successful genes will have is a tendency to postpone the death of their survival machines at least until after reproduction. No doubt some of your cousins and great-uncles died in childhood, but not a single one of your ancestors did. Ancestors just don't die young!

A gene that makes it possessors die is called a lethal gene. A semi-lethal gene has some debilitating effect, such that it makes death from other causes more probable. Any gene exerts its maximum effect on bodies at some particular stage of life, and lethals and semi-lethals are not exceptions. Most genes exert their influence during foetal life, others during childhood, other during young adulthood, others in middle age, and yet others in old age. (Reflect that a

caterpillar and the butterfly it turns into have exactly the same set of genes.) Obviously lethal genes will tend to be removed from the gene pool. But equally obviously a late-acting lethal will be more stable in the gene pool than an early-acting lethal. A gene that is lethal in an older body may still be successful in the gene pool, provided its lethal effect does not show itself until after the body has had time to do at least some reproducing. For instance, a gene that made old bodies develop cancer could be passed on to numerous offspring because the individuals would reproduce before they got cancer. On the other hand, a gene that made young adult bodies develop cancer would not be passed on to very many offspring, and a gene that made young children develop fatal cancer would not be passed on to any offspring at all. According to this theory then, senile decay is simply a by-product of the accumulation in the gene pool of late-acting lethal and semi-lethal genes, which have been allowed to slip through the net of natural selection simply because they are late-acting.

The aspect that Medawar himself emphasizes is that selection will favour genes that have the effect of postponing the operation of other, lethal genes, and it will also favour genes that have the effect of hastening the effect of good genes. It may be that a great deal of evolution consists of genetically-controlled changes in the time of onset of gene activity.

It is important to notice that this theory does not need to make any prior assumptions about reproduction occurring only at certain ages. Taking as a starting assumption that all individuals were equally likely to have a child at any age, the Medawar theory would quickly predict the accumulation in the gene pool of late-acting deleterious genes, and the tendency to reproduce less in old age would follow as a secondary consequence.

As an aside, one of the good features of this theory is that it leads us to some rather interesting speculations. For instance it follows from it that if we wanted to increase the human life span, there are two general ways in which we could do it. Firstly, we could ban reproduction before a certain age, say forty. After some centuries of this the minimum age limit would be raised to fifty, and so on. It is conceivable that human longevity could be pushed up to several centuries by this means. I cannot imagine that anyone would seriously want to institute such a policy.

Secondly we could try to 'fool' genes into thinking that the body they are sitting in is younger than it really is. In practice this would

mean identifying changes in the internal chemical environment of a body that take place during ageing. Any of these could be the 'cues' that 'turn on' late-acting lethal genes. By simulating the superficial chemical properties of a young body it might be possible to prevent the turning on of late-acting deleterious genes. The interesting point is that chemical signals of old age need not in any normal sense be deleterious in themselves. For instance, suppose that it incidentally happens to be a fact that a substance *S* is more concentrated in the bodies of old individuals than of young individuals. *S* in itself might be quite harmless, perhaps some substance in the food which accumulates in the body over time. But automatically, any gene that just happened to exert a deleterious effect in the presence of *S*, but which otherwise had a good effect, would be positively selected in the gene pool, and would in effect *be* a gene 'for' dying of old age. The cure would simply be to remove *S* from the body.

What is revolutionary about this idea is that *S* itself is only a 'label' for old age. Any doctor who noticed that high concentrations of *S* tended to lead to death, would probably think of *S* as a kind of poison, and would rack his brains to find a direct causal link between *S* and bodily malfunctioning. But in the case of our hypothetical example, he might be wasting his time!

There might also be a substance *Y*, a 'label' for youth in the sense that it was more concentrated in young bodies than in old ones. Once again, genes might be selected that would have good effects in the presence of *Y*, but which would be deleterious in its absence. Without having any way of knowing what *S* or *Y* are—there could be many such substances—we can simply make the general prediction that the more you can simulate or mimic the properties of a young body in an old one, however superficial these properties may seem, the longer should that old body live.

I must emphasize that these are just speculations based on the Medawar theory. Although there is a sense in which the Medawar theory logically must have some truth in it, this does not mean necessarily that it is the right explanation for any given practical example of senile decay. What matters for present purposes is that the gene-selection view of evolution has no difficulty in accounting for the tendency of individuals to die when they get old. The assumption of individual mortality, which lay at the heart of our argument in this chapter, is justifiable within the framework of the theory.

The other assumption I have glossed over, that of the existence of sexual reproduction and crossing-over, is more difficult to justify. Crossing-over does not always have to happen. Male fruit-flies do not do it. There is a gene that has the effect of suppressing crossing-over in females as well. If we were to breed a population of flies in which this gene was universal, the *chromosome* in a 'chromosome pool' would become the basic indivisible unit of natural selection. In fact, if we followed our definition to its logical conclusion, a whole chromosome would have to be regarded as one 'gene'.

Then again, alternatives to sex do exist. Female greenflies can bear live, fatherless, female offspring, each one containing all the genes of its mother. (Incidentally, an embryo in her mother's 'womb' may have an even smaller embryo inside her own womb. So a greenfly female may give birth to a daughter and a grand-daughter simultaneously, both of them being equivalent to her own identical twins.) Many plants propagate vegetatively by sending out suckers. In this case we might prefer to speak of *growth* rather than of reproduction; but then, if you think about it, there is rather little distinction between growth and non-sexual reproduction anyway, since both occur by simple mitotic cell division. Sometimes the plants produced by vegetative reproduction become detached from the 'parent'. In other cases, for instance elm trees, the connecting suckers remain intact. In fact an entire elm wood might be regarded as a single individual.

So, the question is: if greenflies and elm trees don't do it, why do the rest of us go to such lengths to mix our genes up with somebody else's before we make a baby? It does seem an odd way to proceed. Why did sex, that bizarre perversion of straightforward replication, ever arise in the first place? What is the good of sex?*

This is an extremely difficult question for the evolutionist to answer. Most serious attempts to answer it involve sophisticated mathematical reasoning. I am frankly going to evade it except to say one thing. This is that at least some of the difficulty that theorists have with explaining the evolution of sex results from the fact that they habitually think of the individual as trying to maximize the number of his genes that survive. In these terms, sex appears paradoxical because it is an 'inefficient' way for an individual to propagate her genes: each child has only 50 per cent of the individual's genes, the other 50 per cent being provided by the sexual partner. If only, like a greenfly, she would bud-off children who were

exact replicas of herself, she would pass 100 per cent of her genes on to the next generation in the body of every child. This apparent paradox has driven some theorists to embrace group-selectionism, since it is relatively easy to think of group-level advantages for sex. As W. F. Bodmer has succinctly put it, sex 'facilitates the accumulation in a single individual of advantageous mutations which arose separately in different individuals.'

But the paradox seems less paradoxical if we follow the argument of this book, and treat the individual as a survival machine built by a short-lived confederation of long-lived genes. 'Efficiency' from the whole individual's point of view is then seen to be irrelevant. Sexuality versus non-sexuality will be regarded as an attribute under single-gene control, just like blue eyes versus brown eyes. A gene 'for' sexuality manipulates all the other genes for its own selfish ends. So does a gene for crossing-over. There are even genes—called mutators—that manipulate the rates of copying-errors in other genes. By definition, a copying error is to the disadvantage of the gene which is miscopied. But if it is to the advantage of the selfish mutator gene that induces it, the mutator can spread through the gene pool. Similarly, if crossing-over benefits a gene for crossing-over, that is a sufficient explanation for the existence of crossing-over. And if sexual, as opposed to non-sexual, reproduction benefits a gene for sexual reproduction, that is a sufficient explanation for the existence of sexual reproduction. Whether or not it benefits all the rest of an individual's genes is comparatively irrelevant. Seen from the selfish gene's point of view, sex is not so bizarre after all.

This comes perilously close to being a circular argument, since the existence of sexuality is a precondition for the whole chain of reasoning that leads to the gene being regarded as the unit of selection. I believe there are ways of escaping from the circularity, but this book is not the place to pursue the question. Sex exists. That much is true. It is a consequence of sex and crossing-over that the small genetic unit or gene can be regarded as the nearest thing we have to a fundamental, independent agent of evolution.

Sex is not the only apparent paradox that becomes less puzzling the moment we learn to think in selfish gene terms. For instance, it appears that the amount of DNA in organisms is more than is strictly necessary for building them: a large fraction of the DNA is never translated into protein. From the point of view of the individual organism this seems paradoxical. If the 'purpose' of DNA is to

supervise the building of bodies, it is surprising to find a large quantity of DNA which does no such thing. Biologists are racking their brains trying to think what useful task this apparently surplus DNA is doing. But from the point of view of the selfish genes themselves, there is no paradox. The true 'purpose' of DNA is to survive, no more and no less. The simplest way to explain the surplus DNA is to suppose that it is a parasite, or at best a harmless but useless passenger, hitching a ride in the survival machines created by the other DNA.*

Some people object to what they see as an excessively gene-centred view of evolution. After all, they argue, it is whole individuals with all their genes who actually live or die. I hope I have said enough in this chapter to show that there is really no disagreement here. Just as whole boats win or lose races, it is indeed individuals who live or die, and the *immediate* manifestation of natural selection is nearly always at the individual level. But the long-term consequences of non-random individual death and reproductive success are manifested in the form of changing gene frequencies in the gene pool. With reservations, the gene pool plays the same role for the modern replicators as the primeval soup did for the original ones. Sex and chromosomal crossing-over have the effect of preserving the liquidity of the modern equivalent of the soup. Because of sex and crossing-over the gene pool is kept well stirred, and the genes partially shuffled. Evolution is the process by which some genes become more numerous and others less numerous in the gene pool. It is good to get into the habit, whenever we are trying to explain the evolution of some characteristic, such as altruistic behaviour, of asking ourselves simply: 'what effect will this characteristic have on frequencies of genes in the gene pool?' At times, gene language gets a bit tedious, and for brevity and vividness we shall lapse into metaphor. But we shall always keep a sceptical eye on our metaphors, to make sure they can be translated back into gene language if necessary.

As far as the gene is concerned, the gene pool is just the new sort of soup where it makes its living. All that has changed is that nowadays it makes its living by cooperating with successive groups of companions drawn from the gene pool in building one mortal survival machine after another. It is to survival machines themselves, and the sense in which genes may be said to control their behaviour, that we turn in the next chapter.