criss-crossed with causal arrows joining genes to phenotypic effects, far and near.

It is an additional fact, too important in practice to be called incidental but not necessary enough in theory to be called inevitable, that these causal arrows have become bundled up. Replicators are no longer peppered freely through the sea; they are packaged in huge colonies—individual bodies. And phenotypic consequences, instead of being evenly distributed throughout the world, have in many cases congealed into those same bodies. But the individual body, so familiar to us on our planet, did not have to exist. The only kind of entity that has to exist in order for life to arise, anywhere in the universe, is the immortal replicator.

EPILOGUE TO 40TH ANNIVERSARY EDITION

Scientists, unlike politicians, can take pleasure in being wrong. A politician who changes his mind is accused of 'flip-flopping'. Tony Blair boasted that he had 'not got a reverse gear'. Scientists on the whole prefer to see their ideas vindicated, but an occasional reversal gains respect, especially when graciously acknowledged. I have never heard of a scientist being maligned as a flip-flopper.

In some ways I would quite like to find ways to recant the central message of *The Selfish Gene*. So many exciting things are fast happening in the world of genomics, it would seem almost inevitable—even tantalizing—that a book with the word 'gene' in the title would, forty years on, need drastic revision if not outright discarding. This might indeed be so, were it not that 'gene' in this book is used in a special sense, tailored to evolution rather than embryology. My definition is that of George C. Williams, one of the acknowledged heroes of the book, now lost to us along with John Maynard Smith and Bill Hamilton: 'A gene is defined as any portion of chromosomal material that potentially lasts for enough generations to serve as a unit of natural selection.' I pushed it to a somewhat facetious conclusion: 'To be strict, this book should be called … *The slightly selfish big bit of chromosome and the even more selfish little bit of chromosome.*' As opposed to the embryologist's

concern with how genes affect phenotypes, we have here the neo-Darwinist's concern with changes in frequencies of entities in populations. Those entities are genes in the Williams sense (Williams later called that sense the 'codex'). Genes can be counted and their frequency is the measure of their success. One of the central messages of this book is that the individual organism doesn't have this property. An organism has a frequency of one, and therefore cannot 'serve as a unit of natural selection'. Not in the same sense of *replicator* anyway. If the organism is a unit of natural selection, it is in the quite different sense of gene 'vehicle'. The measure of its success is the frequency of its genes in future generations, and the quantity it strives to maximize is what Hamilton defined as 'inclusive fitness'.

A gene achieves its numerical success in the population by virtue of its (phenotypic) effects on individual bodies. A successful gene is represented in many bodies over a long period of time. It helps those bodies to survive long enough to reproduce in the environment. But the environment means not just the external environment of the body—trees, water, predators, etc.—but also the internal environment, and especially the other genes with which the selfish gene shares a succession of bodies through the population and down the generations. It follows that natural selection favours genes that flourish in the company of other genes in the breeding population. Genes are indeed 'selfish' in the sense promoted in this book. They are also *cooperative* with other genes with which they share, not just the present particular body, but bodies in general, generated by the species' gene pool. A sexually reproducing population is a cartel of mutually compatible, cooperating genes: cooperating today because they have flourished by cooperating through many generations of similar bodies in the ancestral past. The important point to understand

(it is much misunderstood) is that the cooperativeness is favoured, not because a group of genes is naturally selected as a whole, but because individual genes are separately selected against the background of the other genes likely to be met in a body, and this means the other genes in the species' gene pool. The pool, that is, from which every individual of a sexually reproducing species draws its genes as a sample. The genes of the species (but not other species) are continually meeting each other—and cooperating with each other—in a succession of bodies.

We still don't really understand what drove the origin of sexual reproduction. But a consequence of sexual reproduction was the invention of *the species* as the habitat of cooperating cartels of mutually compatible genes. As explained in the chapter called 'The long reach of the gene', the key to the cooperation is that, in every generation, all the genes in a body share the same 'bottlenecked' exit route to the future: the sperms or eggs in which they aspire to sail into the next generation. *The Cooperative Gene* would have been an equally appropriate title for this book, and the book itself would not have changed at all. I suspect that a whole lot of mistaken criticisms could have been avoided.

Another good title would have been *The Immortal Gene*. As well as being more poetic than 'selfish', 'immortal' captures a key part of the book's argument. The high fidelity of DNA copying—mutations are rare—is essential to evolution by natural selection. High fidelity means that genes, in the form of exact informational copies, can survive for millions of years. Successful ones, that is. Unsuccessful ones, by definition, don't. The difference wouldn't be significant if the potential lifespan of a piece of genetic information was short anyway. To look at it another way, every living individual has been built, during its embryonic development, by genes which can trace their ancestry through a very large number of generations, in a very large number of

individuals. Living animals have inherited the genes that helped huge numbers of ancestors to survive. That is why living animals have what it takes to survive—and reproduce. The details of what it takes vary from species to species—predator or prey, parasite or host, adapted to water or land, underground or forest canopy—but the general rule remains.

A central point of the book is the one developed by my friend the great Bill Hamilton, whose death I still mourn. Animals are expected to look after not only their own children but other genetic relatives. The simple way to express it, and the one that I favour, is 'Hamilton's Rule': a gene for altruism will spread if the cost to the altruist, *C*, is less than the value, *B*, to the beneficiary, devalued by the coefficient of relatedness, *r*, between them. *r* is a proportion between 0 and 1. It has the value 1 for identical twins; 0.5 for offspring and full siblings; 0.25 for grandchildren, half-siblings, and nieces; 0.125 for first cousins. But when is it zero? What is the meaning of zero on this scale? This is harder to say, but it is important and it was not fully spelled out in the first edition of *The Selfish Gene*. Zero does *not* mean that the two individuals share no genes in common. All humans share more than 99 percent of our genes, more than 90 percent with a mouse, and three-quarters of our genes with a fish. These high percentages have confused many people into misunderstanding kin selection, including some distinguished scientists. But those figures are not what is meant by r. Where *r* is 0.5 for my brother (say), it is zero for a *random member of the background population with whom I might be competing*. For purposes of theorizing about the evolution of altruism, *r* between first cousins is 0.125 only when compared to the reference background population ($r = 0$), which is the rest of the population to whom altruism potentially might have been shown: competitors for food and space, fellow travellers through time in the environment of the species. The 0.5 (0.125, etc.) refers

to the *additional* relatedness *over and above* the background population, whose relatedness approaches zero.

Genes in the Williams sense are things you can count as the generations go by, and it doesn't matter what their molecular nature is; it doesn't matter, for instance, that they are split up into a series of 'exons' (expressed) separated by mostly inert 'introns' (ignored by the translation machinery). Molecular genomics is a fascinating subject but it doesn't heavily impinge on the 'gene's eye view' of evolution which is the central theme of the book. To put the point another way, *The Selfish Gene* is quite likely a valid account of life on other planets even if the genes on those other planets have no connection with DNA. Nevertheless, there are ways in which the details of modern molecular genetics, the detailed study of DNA, can be gathered into the gene's eye fold and it turns out that they vindicate that view of life rather than casting doubt on it. I'll come on to this after what may seem like a radical change of subject, beginning with a specific question which obviously stands for any number of similar questions.

How closely related are you to Queen Elizabeth II? As it happens, I know I'm her 15th cousin twice removed. Our common ancestor is Richard Plantagenet, 3rd Duke of York (1412–1460). One of Richard's sons was King Edward IV, from whom Queen Elizabeth is descended. Another son was George, Duke of Clarence, from whom I am descended (allegedly drowned in a butt of Malmsey wine). You may not know it but you are very probably closer to the Queen than 15th cousin and so am I and so is the postman. There are so many different ways of being somebody's distant cousin, and we are all related to each other in many of those ways. I know that I am my wife's twelfth cousin twice removed (common ancestor George Hastings, 1st Earl of Huntingdon, 1488–1544). But it is highly probable that I am a closer cousin to her in various unknown ways (various pathways

through our respective ancestries) and it is absolutely certain I am also her more distant cousin in many more ways. We all are. You and the Queen might simultaneously be ninth cousins six times removed, and twentieth cousins 4 times removed, and thirtieth cousins 8 times removed. All of us, regardless of where in the world we live, are not only cousins of each other. We are cousins in hundreds of different ways. This is just another way of saying we are all members of the background population among whom r, the coefficient of relatedness, approaches zero. I could calculate r between me and the Queen using the one pathway for which records exist, but it would, as the definition requires, be so close to zero as to make no difference.

The reason for all that bewildering multiplicity of cousinship is sex. We have two parents, four grandparents, eight great grandparents, and so on, up to astronomical numbers. If you go on multiplying by two back to the time of William the Conqueror, the number of your ancestors (and mine, and the Queen's and the postman's) would be at least a billion, which is more than the world population at the time. That calculation alone proves that, wherever you come from, we share many of our ancestors (ultimately all if you go sufficiently far back) and are cousins of each other many times over.

All that complexity disappears if you look at cousinship from the gene's point of view (the point of view advocated, in different ways, throughout this book) as opposed to the individual organism's point of view (as has been conventional among biologists). Stop asking: What kind of cousin am I to my wife (the postman, the Queen)? Instead, ask the question from the point of view of a single gene, say my gene for blue eyes: What relation is my blue eye gene to the postman's blue eye gene? Polymorphisms such as the ABO blood groups go way back in history, and are shared by other apes and even monkeys. The A gene in a human sees the

equivalent gene in a chimp as a closer cousin than the B gene in a human. As for the SRY gene on the Y chromosome, which determines maleness, my SRY gene 'looks upon' the SRY gene of a kangaroo as its kissing cousin.

Or we can look at relatedness from a mitochondrion's point of view. Mitochondria are tiny bodies teeming in all our cells, vital to our survival. They reproduce asexually and retain the remnants of their own genomes (they are remotely descended from free-living bacteria). By the Williams definition, a mitochondrial genome can be thought of as a single 'gene'. We get our mitochondria from our mothers only. So if we were now to ask how close is the cousinship of your mitochondria to the Queen's mitochondria there is a single answer. We may not know what that answer is, but we do know that her mitochondria and yours are cousins in only one way, not hundreds of ways as is the case from the point of view of the body as a whole. Trace your ancestry back through the generations, but always only through the maternal line and you follow a single narrow (mitochondrial) thread, as opposed to the ever branching thread of 'whole organism pedigrees'. Do the same for the Queen, following her narrow maternal thread back through the generations. Sooner or later the two threads will meet and now, by simply counting generations along the two threads, you can easily calculate your mitochondrial cousinship to the Queen.

What you can do for mitochondria, you can in principle do for any particular gene, and this illustrates the difference between a gene's point of view and an organism's point of view. From a whole organism point of view you have two parents, four grandparents, eight great grandparents, etc. But, like a mitochondrion, each gene has only one parent, one grandparent, one great grandparent, etc. I have one gene for blue eyes and the Queen has two. In principle we could trace the generations back and discover

the cousinship between my blue eye gene and each of the Queen's two. The common ancestor of our two genes is called the 'coalescence point'. Coalescence analysis has become a flourishing branch of genetics and very fascinating it is. Can you see how congenial it is to the 'gene's eye view' that this whole book espouses? We are not talking about altruism any more. The gene's eye view is flexing its muscles in other domains, in this case looking back at ancestry.

You can even investigate the coalescence point between two alleles in one individual body. Prince Charles has blue eyes and we can assume that he has a pair of blue eye alleles opposite each other on Chromosome 15. How closely related to each other are Prince Charles' two blue eye genes, one from his father, one from his mother? In this case we know one possible answer, only because royal pedigrees are documented in ways that most of our pedigrees are not. Queen Victoria had blue eyes and Prince Charles is descended from Victoria in two ways: via King Edward VII on his mother's side; and via Princess Alice of Hesse on his father's side. There's a 50 percent probability that one of Victoria's blue eye genes peeled off two copies of itself, one of which went to her son, Edward VII, and the other to her daughter Princess Alice. Further copies of these two sibling genes could easily have passed down the generations to Queen Elizabeth II on one side and Prince Philip on the other, whence they were reunited in Prince Charles. This would mean the 'coalescence' point of Charles's two genes was Victoria. We do not—cannot—know whether this actually is true for Charles' blue eye genes. But statistically it has to be true that many of his pairs of genes coalesce back in Victoria. And the same applies to pairs of your genes, and pairs of my genes. Even though we may not have Prince Charles' well-documented pedigree to consult, any pair of your genes could, in principle, look back at their common ancestor, the

coalescence point at which they were 'peeled off' from the same parent gene.

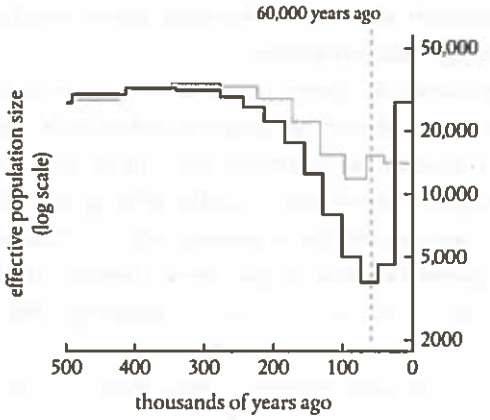
Now, here's something interesting. Although I can't establish the exact coalescence point of any particular allelic pair of my genes, geneticists can in principle take all the pairs of genes from any one individual and, by considering all possible pathways back through the past (actually not all possible pathways because there are too many, but a statistical sample of them), derive a pattern of coalescences across the whole genome. Heng Li and Richard Durbin of the Sanger Institute in Cambridge realized a remarkable thing: the pattern of coalescences among pairs of genes in the genome of a single individual gives us enough information to reconstruct demographic details about datable moments in the prehistory of an entire species.

In our discussion of coalescence between pairs of genes, one from the father and one from the mother, the word 'gene' means something a bit more fluid than the normal usage of molecular biologists. Indeed, you could say that the coalescence geneticists have reverted to something a bit like my 'slightly selfish big bit of chromosome and even more selfish little bit of chromosome.' Coalescence analysis is studying chunks of DNA which might be larger or even smaller than a molecular biologist's understanding of a single gene but which can still be seen as cousins of each other, having been 'peeled off' from a shared ancestor some definite number of generations ago.

When a gene (in that sense) 'peels off' two copies of itself and gives one to each of two offspring, the descendants of those two copies may, over time, accumulate differences due to mutation. These may be 'under the radar' in the sense that they don't show up as phenotypic differences. The mutated differences between them are proportional to the time that has elapsed since the split, a fact that biologists make good use of, over much greater time

spans, in the so-called 'molecular clock'. Moreover, the pairs of genes whose cousinship we are calculating needn't have the same phenotypic effects as each other. I have one blue eye gene from my father paired with one brown eye gene from my mother. Although these genes are different, even they must have a coalescence somewhere in the past: the moment when a particular gene in a shared ancestor of my two parents peeled off one copy for one child and another copy to its sibling. That coalescence (unlike the two copies of Victoria's blue eye gene) was a long time ago, and the pair of genes has had a long time to accumulate differences, not least the difference in the eye colours that they mediate.

Now, I said that the coalescence pattern within one individual's genome can be used to reconstruct details of demographic prehistory. Any individual's genome can do this. As it happens, I am one of the people in the world who has had his complete genome sequenced. This was for a television programme called *Sex, Death and the Meaning of Life* which I presented on Channel Four in 2012. Yan Wong, my co-author of *The Ancestor's Tale*, from whom I learned everything I know about coalescence theory and much else besides, seized upon this and did the necessary Li/Durbin style calculations using my genome, and my genome alone, to make inferences about human history. He found a large number of coalescences around 60,000 years ago. This suggests that the breeding population in which my ancestors were embedded was small 60,000 years ago. There were few people around, so the chance of a pair of modern genes coalescing in the same ancestor back at that time was high. There were fewer coalescences 300,000 years ago, suggesting that the effective population size was larger. These figures can be plotted as a graph of effective population size against time. Here's the pattern he found, and it is the same pattern as the originators of the technique would expect to find from any European genome.



From R. Dawkins and Y. Wong (2006) *The Ancestor's Tale*, 2nd edition
Image courtesy of Y. Wong

The black line shows the estimates of effective population size at various times in history based upon my genome (coalescences between genes from my father and my mother). It shows that the effective population size among my ancestral population plummeted around 60,000 years ago. The grey line shows the equivalent pattern derived from the genome of a Nigerian man. It also shows a drop in population around the same time, but a less dramatic one. Perhaps whatever calamity caused the drop was less severe in Africa than in Eurasia.

Incidentally Yan was my undergraduate pupil in New College, Oxford, before I started learning more from him than he learns from me. He then became a graduate student of Alan Grafen, whom I had also tutored as an undergraduate, who subsequently became my graduate student and whom I have described as being now my intellectual mentor. So Yan is both my student and my grandstudent—a neat memetic analogue to the point I was making earlier about how we are related in multiple ways—although

the direction of cultural inheritance is more complicated than this simple formulation implies.

To summarize, the gene's eye view of life, the central theme of this book, illuminates not just the evolution of altruism and selfishness, as expounded in previous editions. It also illuminates the deep past, in ways of which I had no inkling when I first wrote *The Selfish Gene* and which are expounded more fully in relevant passages (largely written by Yan, my co-author) of the second edition (2016) of *The Ancestor's Tale*. So powerful is the gene's eye view, the genome of a single individual is sufficient to make quantitatively detailed inferences about historical demography. What else might it be capable of? As foreshadowed by the Nigerian comparison, future analyses of individuals from different parts of the world could give a geographic dimension to these demographic signals from the past.

Might the gene's eye view penetrate the remote past in yet other ways? Several of my books have developed an idea which I called 'The Genetic Book of the Dead'. The gene pool of a species is a mutually supportive cartel of genes that have survived in particular environments of the past, both distant and recent. This makes it a kind of negative imprint of those environments. A sufficiently knowledgeable geneticist should be able to read out, from the genome of an animal, the environments in which its ancestors survived. In principle, the DNA in a mole *Talpa europaea* should be eloquent of an underground world, a world of damp, subterranean darkness, smelling of worms, leaf decay, and beetle larvae. The DNA of a dromedary, *Camelus dromedarius*, if we but knew how to read it, would spell out a coded description of ancient ancestral deserts, dust storms, dunes, and thirst. The DNA of *Tursiops truncatus*, the common bottlenose dolphin, spells, in a language that we may one day decipher, 'open sea, pursue fish fast, avoid killer whales'. But the same dolphin DNA

also contains paragraphs about earlier worlds in which the genes also survived: on land when the ancestors escaped the attentions of tyrannosaurs and allosaurs long enough to breed. Then, before that, parts of the DNA surely spell out descriptions of even older feats of survival, back in the sea, when the ancestors were fish, pursued by sharks and even eurypterids (giant sea scorpions). Active research on 'The Genetic Book of the Dead' lies in the future. Will it colour the epilogue of the fiftieth anniversary edition of *The Selfish Gene?*