The MUSIC of LIFE

Biology Beyond Genes

Denis Noble

The MUSIC *of* LIFE

Biology beyond Genes

Denis Noble

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2006
First published as an Oxford University Press paperback 2008

British Library Cataloguing in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain by
Clays Ltd, St Ives plc

ISBN 978–0–19–922836–2

1 3 5 7 9 10 8 6 4 2

Contents

This page intentionally left blank

Introduction

'What is life?' The question can be interpreted in many ways. One way to approach it is scientifically. Even from this standpoint there can be a variety of answers since contemporary scientists can understand the question rather differently. Moreover, each generation needs to revisit the question almost anew – the advances in biological science are that rapid.

It is only 50 years ago that we first discovered that the genetic material was the molecule called DNA (deoxyribonucleic acid), and that it came in long molecular strands of four similar chemicals, called bases. Now:

- We know that the human genome, the entire DNA of a human, is a sequence of 3 billion pairs of bases – and we have identified each of them.
- We also know how the organisation of these bases serves to enable protein production. For each protein, the genetic material provides something like a template. The structural sequence of the proteins is encoded in the DNA. We know in some detail how this code works.
- For that matter, we also know the sequences and structure of many of the proteins that the DNA encodes.

Biological science has never advanced so rapidly.

How has that changed the way we see life? It has answered many questions and thrown up many more. The answers that we arrive at reflect the process of investigation that we follow. Over the last half century, we have proceeded by breaking living systems down into their smallest components, the individual genes and molecules. Humpty-Dumpty has been smashed into billions of fragments. This is an impressive achievement.

For example, we can now pinpoint a gene mutation whose effects

may 'kick in' during middle age to cause sudden cardiac death. We know nearly all the major steps in this causal chain, though not yet why it kicks in precisely when it does in a given individual. This kind of success is more and more common. Yet, such examples are not appearing with the frequency that optimists predicted when the human genome project was announced. The benefits for healthcare are slow to arrive.

Why is that? People are beginning to understand the reason. It has to do with how the small scale relates to the large. We know a lot about molecular mechanisms. Now the challenge is to extend that knowledge up the scale. How do we use it to throw light on the processes that govern entire living systems? That is not an easy question. Quite soon, as we move from genes to the proteins that they code for, and then on to the interactions between these proteins, the problems become seriously complicated. Yet we need to understand these complexities in order to interpret the molecular and genetic data, and on that basis to talk in a fresh and useful way about larger questions like 'What is life?'

This, then, is the challenge that sequencing the genome has raised. Can we put Humpty-Dumpty back together again? That is where 'systems biology' comes in. This is a new and important dimension of biological science, though it has strong historical roots in classical biology and physiology going back over a century. In recent decades, however, biologists have tended to focus quite narrowly on the individual components of living organisms. What properties does each component have? How does it therefore interact, over the short term, with other components of similar scale? Now, we are ready to ask some bigger questions. These are about systems. At each level of the organism, its various components are embedded in an integrated network or system. Each such system has its own logic. It is not possible to understand that logic merely by investigating the properties of the system's components.

This book is about systems biology. It is also about the preconditions for, and implications of systems biology. It says that at this stage in our exploration of life, we need to be ready for a basic re-think.

Molecular biology requires a certain way of thinking. It is about

the naming and behaviour of the parts. We reduce each whole to its component parts and define them exhaustively. Biologists are now perfectly used to that thinking and the interested lay public has caught up, too. So we are now ready to move on. Systems biology is where we are moving to. Only, it requires a quite different mind-set. It is about putting together rather than taking apart, integration rather than reduction. It starts with what we have learned from the reductionist approach; and then it goes further. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist procedures, but different. This is a major change. It has implications beyond the purely scientific. It means changing our philosophy, in the full sense of the term.

How to provoke such a change? I have chosen to write a polemic. This book is a radical analysis of many of the currently accepted dogmas in biology. It turns some of them upside down. It offers an unashamed defence of the need for a systems level approach. That is not because I am unimpressed with what reductionist molecular biology has achieved. On the contrary, it is because I want to see biological science garner the fruits that the great reductionist drive has put within our grasp.

As I explain in Chapter 5, I started my research career in physiology as a full 'card-carrying' reductionist. I know how successful reductionist science is done and have done much of it myself in my own field. I still use its methods quantitatively in my current research on simulating the organs of the body. And that is how, during the last decade or so, I have come to see the need to redress the balance. If we all keep our noses down to the lower-level grindstone, no-one will see the bigger picture, or realise what is needed if we are to fill it in. Successful integration at the systems level must be built on successful reduction, but reduction alone is far from sufficient.

Like any polemicist, I make free use of metaphor. I also tell some stories. These are intended to be enjoyable – and also to jolt the reader away from many current dogmas.

In 1944, Erwin Schrödinger wrote a remarkable book (Schrödinger 1944). In it, he correctly predicted that the genetic code is an 'aperiodic crystal', that is, a chemical sequence without regular

repetition. Like many scientists at that time, he thought that the code would be found in the proteins rather than in DNA, so what he spoke of was not where he expected it – but it was there nonetheless. Many of his insights match remarkably well with what we have since learned. In just under 100 pages, he shifted the basic paradigms of biology.

This book is of similar length. I first thought to give it the same title: 'What is life?' But I have not been so audacious. Instead, I have chosen a title that reflects the main metaphor of the book: the systems-level view of life can be compared to music. If so, where is the score and who was the composer? A central question, therefore, that recurs throughout the book is 'Where, if anywhere, is the program of life?' The French Nobel Prize-winners Jacques Monod and François Jacob (Monod and Jacob 1961; Jacob 1970), referred to the 'genetic program' (le programme génétique): the idea that the instructions for the development of each living organism lie in its genes. The same idea is conveyed by the popular description of the genome as the 'book of life', a kind of blueprint. The central role of genes as causal agents was also greatly reinforced by popular perceptions of Richard Dawkins' highly influential book *The selfish gene* (Dawkins 1976).

The theme of my book is that there is no such program and that there is no privileged level of causality in biological systems. Chapter 1 lays the groundwork for the rest of the book. It does this first by recasting the genome as a database for the transmission of successful organisms, rather than a program that 'creates' them. The second step is to replace the metaphor of the 'selfish gene' by 'genes as prisoners'. These two radical switches of perception are essential to understanding the rest of the book. While it is necessary to deal with the popular (mis)perceptions of 'genetic programs', 'the book of life', and 'selfish genes', I hasten to acknowledge that the scientists responsible for these fruitful ideas may well not have approved of the way they have been widely interpreted. Richard Dawkins, for example, has also written some of the best critiques of the 'program' idea, and is himself far from being a gene determinist.

The book is organised into ten chapters. Each uses a different musical metaphor for some aspect of the biology of life. We start with

the genome in Chapter 1 and end with the brain in Chapter 9. Chapter 10 stands on its own as a kind of coda.

Acknowledgements

I acknowledge valuable discussions on various aspects of this book with Geoff Bamford, Patrick Bateson, Steven Bergman, Sydney Brenner, Jonathan Cottrell, Christoph Denoth, Dario DiFrancesco, Yung Earm, David Gavaghan, Peter Hacker, Jonathan Hill, Peter Hunter, Otto Hutter, Roger Kayes, Anthony Kenny, Sung-Hee Kim, Junko Kimura, Peter Kohl, Jean-Jacques Kupiec, Ming Lei, Nicholas Leonard, Jie Liu, Denis Loiselle, Latha Menon, Alan Montefiore, Penny Noble, Ray Noble, Susan Noble, Carlos Ojeda, Etienne Roux, Ruth Schachter, Pierre Sonigo, Christine Standing, Richard Vaughan-Jones and Michael Yudkin. Many of these, and the Oxford University Press readers, also criticised early versions of various chapters. I benefited greatly from their feedback, though I am of course responsible for the errors and misunderstandings that remain.

I thank many friends and colleagues in East Asia who introduced me to aspects of their cultures that find expression in Chapters 8 and 10.

An early version of the Silman story in Chapter 1 appeared in French under the title 'Pourquoi il nous faut une théorie biologique' in the online publication *Vivant* in 2004. Parts of Chapter 3 are based on 'Is the genome the book of life?', *Physiology News* (2002), **46**, 18–20. The dialogue in Chapter 9 is based on 'Qualia and private languages', *Physiology News* (2004), **55**, 32–3, while the following story first appeared as 'Biological explanation and intentional behaviour' in *Modelling the mind* (ed. K.A. Mohyeldin Said *et al.*, Clarendon Press, Oxford, 1990), pp. 97–112. Some of the philosophical background was developed in *Goals, no goals and own goals* (ed. A. Montefiore and D. Noble, Unwin Hyman, London, 1989) and in various Novartis Foundation meetings on the nature of biological science.

This page intentionally left blank

1 The CD of Life: the Genome

They [genes] are all in the same boat.
Maynard Smith and Szathmáry 1999

For humans at least, to live is to experience. How can we understand this?

One thing is clear: experience is grounded in matter. The connection is there for us to draw. But drawing it is quite a complicated task. And, alas, complexity is uncomfortable, so we are inclined to ignore it.

That tends to happen, for instance, when we try to tease out the linkage between human experience and material reality. We say, 'It's pretty simple, really'. But it isn't.

Introducing the Silmans

Consider an example. Before writing this page, I relaxed by listening for the first time for a long time to one of my favourite pieces of music: the piano trio in E-flat major by Schubert. I put the CD into my player and lay down on the sofa. As the music entered the slow movement, I cried.

The emotional effect of this piece of music, which I first heard live in a concert, is always very strong. We must all have our favourite pieces that have this effect on us. The effect does not always depend on the music itself. It can also depend on the context, the people we were with, and the significance of the event in our lives.

So, what caused me to cry?

Imagine some space travellers watching this scene. They are creatures from a world in which silicon replaces carbon. So let's call them Silmans. They have some of the characteristics associated in science fiction with 'androids'. They notice the crying. They record the sound waves in the room. As scientists, they trace the sequence of cause and effect, back through the loudspeakers, the amplifiers, the laser disc reader, right down to the CD itself.

One of them does a Silman version of 'Eureka!' 'I've found it', he says, as he explains to his colleagues that the whole effect is caused by some highly specific digital information on the CD. Another of the Silmans is nevertheless sceptical. 'How', he says, 'could just a bunch of numbers have this effect?'

The discoverer counters the scepticism by pointing out that this is the lowest level of the chain of cause and effect. Without the digital information, there would be no music, no emotion. Moreover, if you play around with that information, 'mutate' it as it were, by playing it too fast or too slow, or playing it backwards, transposing sections, or even transposing bits from another CD, then the person in the room no longer cries. In fact he may angrily turn the machine off and even throw the disc away.

There is an inevitable and mechanical chain of cause and effect here. Any experiment the Silmans might do would reinforce the one-way nature of this chain. Different amplifiers, speakers, and other gadgetry can replace everything except the highly specific digital information on the CD. Surely, then, they conclude that this is the cause of me crying.

Of course, we know better. We would say that the causes of my crying include:

- Schubert, because he wrote the music;
- the piano trio, because they played it with such heart-tugging inspiration;
- and the beautiful context in which I first heard the music and first cried as a result of it. This, we would say, is in my memory and forms the emotional context.

We would say that the digital information on the CD is just a way of

capturing the moment, as accurately as possible, and making it possible for me to recreate, partially at least, the original moment. We know also that the information could be coded in many different ways, including analogue encoding in the form of a vinyl disc. It is just a database that enables the music to be stored and recreated.

In short, we would have no difficulty at all in laughing at the stupidity of our Silman visitors from another planet. They saw a simple explanation, we would say, and grabbed at it. How stupid! Well, we should be careful whom we laugh at. For we, too, get trapped in simplistic explanations.

DNA-mania

Indeed, there is a popular dogma that is reinforced daily in the media—and, it must be said, by many scientists—that rests on a crude mistake, just like the Silmans'. André Pichot[1] has called this DNA-mania. It is the delusion that the DNA code 'causes' life in much the same way as the CD 'caused' my experience of the Schubert piano trio.

The analogy is obvious. The human genome is in some ways a little like a CD. It carries digital information. Let's quickly summarise how. The genome is all the chromosomes in a cell. A chromosome is a long DNA molecule and some associated proteins. It is conventionally divided into genes. A gene is a section of DNA that is used in producing a particular protein.

DNA is composed of four chemicals (nucleotides), generally referred to by the letters A, T, G, and C.[2] There are two strands of DNA in each chromosome, wrapped around each other in a double helix. It was the discovery of this double helical structure that formed Watson and Crick's Nobel Prize-winning work in 1953. The nucleotides in one strand always lie opposite those in the other according to the rule: A goes opposite T, G opposite C. Two such

[1] A French philosopher and historian of science, author of *Histoire de la notion de gène* (Pichot 1999).

[2] Adenine, thymine, guanine, and cytosine.

complementary nucleotides make a base pair. The genome is 3 billion base pairs long. These form 20 000–30 000 genes.

In each gene, the chemicals are arranged in specific ways to facilitate the production of specific proteins. Every time a protein is needed, the appropriate chemical 'code' is 'read off' the gene; this gives the pattern of chemical elements that will make that protein what it is. Our genes encode the sequences of the 100 000 or so proteins that make up the human body. No protein is made that is not coded for by a gene. So the genome *is* important. After all, proteins are crucial for life.

A living cell is a continuing, action-packed drama. Molecules interact and change. One change triggers another, and so on and on. Complex chains of molecular interaction happen again and again. We call them 'pathways'. There are cell cycle pathways, which correspond to the cell 'ticking over'. There are developmental pathways, because cells grow, divide, and form more cells. There are all sorts of regulatory pathways. And proteins form the backbone of all these biochemical pathways.

Cells organise into tissues, such as skin, bone, muscle, to form organs such as the heart and kidneys, and finally, all these, together with the immune and hormonal systems, form the organism, the whole animal. This operates in many different ways, at various levels of organisation. And all of this 'function', as biologists say, involves proteins.

The causality seems to be entirely one-way. The DNA causes the proteins, the proteins cause the cells, and so on. The organism itself is just what shows on the outside; what is really happening, the inside story, is that the information coded in the genes is being expressed. In biologist-speak, the phenotype is 'created by' the genotype. The story is seductive.

We have fitted ourselves out with a magnificent set of blinkers. We have rendered ourselves incapable of looking at the relationships between the genetic code and living systems in any other way.

This chapter asks why.

- Why are we so fond of the gene-centred view? We can explore

this question by examining a classic and very popular statement of this view—Dawkins' 1976 description of the 'selfish gene'.

- How did so many people come to interpret this view as genetic determinism? That question is particularly important since, as I will show, it is not that of Dawkins himself. Let us then explore the historical context, out of which DNA-mania has developed.

We start with the reductionist causal chain. This is the 'inside story' that we have just discussed. Schematically, it looks like this:

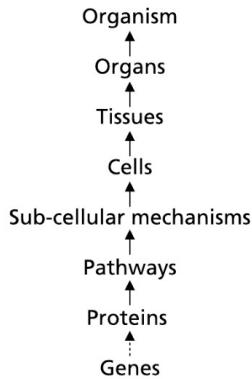

FIG 1. The reductionist causal chain.

The chain runs upwards. It is a 'one-way' system, from the genes to the organism. The idea is that, if we knew all about the lowest-level elements, genes and proteins, then everything about the organism would be clear to us. We could work out what happens at the higher levels, and explain it completely, in terms of our low-level knowledge. We could reconstruct the whole organism from the bottom up.

The first step in the chain is fainter than the others because it represents a rather different causal relationship. At each stage above this one, we are talking about physical causes—how one chemical reaction leads to another. But at the first stage something different is happening, over and above the physical causation of the chemical reactions involved. It is generally described as the reading of a code. There is transcription and translation of the code. This code is

sometimes called the blueprint of life, or the program of life, following Monod and Jacob's colourful idea of 'le programme génétique' (Monod and Jacob 1961; Jacob 1970).

So much for the diagram. The problem with it is that it shows only half the story. When we get to Chapter 4, we shall see how much it misses out. But for the moment let us assume it is as comprehensive as it is supposed to be.

On that assumption, then, let us ask: does the causal mechanism work in the way that is represented here? By no means!

Problems with genetic determinism

Genes are coded as DNA sequences. It is these sequences that are replicated and passed on to future generations. So biologists also call genes replicators. Gene determinism somehow sees them as causal agents. How can that be? After all, what does DNA *do*? As biological molecules go, not much. The real players in the action of life are the proteins. They are the really active molecules. They indulge most in the biochemical processes necessary for life to occur. DNA is in comparison rather passive.

Proteins are produced in tiny factories inside the cells of the body. Biologists call them ribosomes. These factories get going when they receive a message that 'tells' them to make a certain protein. Each such message is generated using DNA. A DNA sequence that corresponds to the relevant protein sequence is copied onto another molecule, appropriately called a 'messenger', which transmits a form of the sequence to the ribosomes. The messenger molecules, called messenger RNA (ribonucleic acid), are another kind of nucleic acid sequence. The DNA sequences are therefore a kind of template, a specific sequence of nucleotides that can be transcribed to produce the message that is then translated into an amino-acid sequence when the protein is made. (Amino acids are the units of which protein is composed, just as nucleotides are the units of which DNA is composed.)

That process is called 'gene expression'. This terminology gives the impression that the whole process is implicit in the gene, or at least in

the information that the gene holds, which simply needs to be 'expressed'.

But it is a little odd to say, as we often do, that the DNA sequence 'determines' the protein. In fact, the DNA just sits there, and occasionally the cell reads off from it a sequence that it needs, in order to get some protein produced. This looks very much like my hi-fi equipment reading the digital information on a CD to generate the real 'action': the music. So the first step in the reductionist chain of cause and effect is not a simple causal event at all. When a sequence is read off, that is an important event, which initiates a whole series of subsequent events. These are physical events. True. But it is the process of reading that matters, as well as the object that is read.

This process involves certain systems of proteins. If we wish to identify an agent of the action, it must be those systems. They 'read' the DNA code. DNA does nothing outside the context of a cell[3] containing these protein systems, just as the CD can do nothing without the CD reader. So, we have the paradox that proteins are required for the machinery to read the code to produce the proteins. I will return to this paradox in later chapters.

But is this just a technicality? Whether we start the causal chain from genes or proteins doesn't matter much, perhaps. Don't we just have to adjust our story slightly and say that the genetic code lies in the protein sequences? That might be a reasonable way of looking at the matter, except that it assumes that each gene codes directly for a single protein, that is, that the two sequences, in the DNA and the protein, are straightforwardly identical. But they are not.

In higher animals, the bits of DNA code that we lump together and call collectively a 'gene' are not always continuous. In many, perhaps most cases, they are broken up into segments. These segments, called 'exons', are separated by non-coding stretches of DNA, called 'introns'. The exon codes can be combined in various orders to produce a full protein code. In ways we do not yet understand, the DNA threads are folded into a three-dimensional form in the nucleus of

[3] Viruses are not an exception to this rule. They need to enter a cell to use its machinery to reproduce. Outside a cell they cannot reproduce.

each cell. They cannot exist as a straight thread since each cell contains two metres of DNA, which is around 100 000 times longer than most cells. The way the threads are folded inside the cell may make the reading of certain sequences easier than others.

There can therefore be many different ways of reading the separate exons and joining them up. Technically, there are often many 'splice variants' of a gene, which can therefore code for a set of different proteins. These splice variants are the different ways of reading the separate exons and joining them up (Black 2000). Thus, if a gene consists of three exons, *a*, *b*, and *c*, it could be read as the forms *a*, *b*, *c*, *ab*, *bc*, *ac*, *abc*, and perhaps even as *cba*, *ca*, *ba*, each of which would code for a different protein. At present, we do not know the rules for which combinations are possible and used in coding for proteins.

Consider the gene called *Dscam* in the fruit fly *Drosophila*. It has 110 introns and therefore tens of thousands of possible splice variants (Celotto and Graveley 2001). Moreover the *Dscam* gene does not always operate in the same way. It changes its role with the life cycle of *Drosophila*. At any one stage, some of the theoretically possible splice variants will work and others will not. At earlier and later stages, the picture will be different.

To an extent, it depends on the cellular environment. For instance, there are proteins that affect the transcription of DNA sequences. Some activate transcription; some inhibit it; they interact in complex ways. At the same time, there are features in the DNA code itself that influence whether a particular variant can be expressed. Within the DNA sequences of a given gene we find promoter elements and enhancer sequences. So the regulation of gene expression, as we say, involves a multiplicity of factors which operate and inter-operate in subtle ways.

There is regulation of how the code is read off from the gene in order to form the protein (transcription), and regulation of what happens after transcription. These are all complicated processes subject to many other influences than the DNA code itself.

What this means is that there are many different ways to read a genome. So my analogy with a CD is limited. When you put a CD into your hi-fi, there is only one way you are going to get music out

of each track. This process involves a single, one-way read-out. When it comes to genes, by contrast, we have flexible, combinatorial read-out. A clever CD reader can also do this to some extent. We can program our hi-fi equipment to play the music tracks in flexible orders depending on how the recording has been broken up to form the individual tracks. The difference is that the genome is broken up to an unimaginably greater degree, the consequences of which we will explore in Chapter 2.

Included in this flexibility are many back-up processes. So it is possible to correct for errors and failures at the genome level. Indeed, it can happen that an important gene is completely knocked out, and the organism still manages to get by. If Plan A does not work, Plan B clicks in—the cell is still able to form proteins which function in place of those the non-functioning gene was originally used for.

To these multiple influences at the bottom level of protein production we must add an important higher-level complexity. This is that there is no one-to-one correspondence between genes and biological functions. Strictly speaking, therefore, to speak of a gene as the 'gene for x' is *always* incorrect. Many gene products, the proteins, must act together to generate biological functions at a high level. If we must use the expression 'gene for x' then we should at least add the plural and speak of the 'genes for x'.

Even this way of speaking is, however, seriously misleading. Not only do many genes co-operate in coding for the proteins that interact to produce any given biological function, each gene may also play a role in many different functions, which makes it difficult to label genes with functions.

I am talking here about higher-level function in organisms. Let's think of some of these functions. The pacemaker rhythm of the heart would be one. Another would be the secretion of insulin by the pancreas. And then we could take the transmission of impulses in the brain. Then let's think of the lower-level biological processes involved in those functions. For instance there is the process whereby calcium ions get pushed out of cells. Certain identifiable proteins combine to produce this effect. They are important because calcium is used as a controller of many processes in cells and organs.

This process of moving calcium ions goes on in all sorts of ways and contexts. For instance, it is involved in all three of the functions we have identified, and many more besides. In fact it is hard to think of a single higher-level function that does not involve these calcium-pushing proteins—and so, implicitly, the genes that code for them. I could repeat the same story for many other processes in cells. Many of the lower-level processes are used again and again in many different functions. High-level functions are therefore like a game of recombinations.

Suppose we sat down to identify the role or roles that those genes play in high-level functions. We would end up with a list that went on almost forever. That is what happens when we start to study how biological function emerges. We get involved in an almost endless game of recombinations. So, while it is relatively easy to label genes with low-level functions, which proteins they code for, it is much more difficult to label genes with high-level functions.

We need a manual that lists all the functions a gene is involved in, and how it contributes to each. Nature does not provide one. We have to work this stuff out for ourselves. That is the research project we call gene ontology. And to get anywhere with that project, we have to look beyond genes and proteins. We need to study the higher-level functions.

This is my primary reason for opposing the otherwise colourful metaphor describing the genome as 'the book of life' (Chapter 3). A book may describe, explain, illustrate, and may do many other things, but if we opened it up to find just strings of numbers, like the machine code of a computer program, we would surely ask where is the book itself; we would say that we had only been given a database. We could, perhaps, use another, interpretative program to generate a 'book' from it; but, until then, all we would have would be a mass of ciphers.

My central argument will be that the book of life is life itself. It cannot be reduced to just one of its databases. For let's be clear that the genome is only *one* of the databases. Function in biological systems depends also on important properties of matter that are not specified by genes. We will return to this aspect in Chapter 3.

Origin of the appeal of genetic determinism

There is work to do before we can start to make much sense of the genetic information we have discovered. The problems are immense, as we shall see in Chapter 2. Indeed, we must wonder how long it will take us to overcome them.

Why, then, has the genetic determinist agenda had such a wide and fashionable appeal? How does it come to dominate the way in which public debate on genes takes place, with 'genes for this' and 'genes for that' appearing with regular frequency, implying that it is only a matter of time before we find the genes for everything? This is where we need to look at the history of the development of ideas on genetics and on biology as a science.

There is an interesting contrast between the ways in which these ideas have developed in French-speaking and English-speaking countries. I will refer to the debate in French-speaking countries later in this book. In the Anglo-Saxon world the debate has been dominated by arguments between the gene-centred views of people like Richard Dawkins (1976) and the multi-level selection views of people like Stephen Jay Gould (2002).

The gene-centred view, the 'selfish gene' view, is a metaphorical polemic: the invention of a colourful metaphor to interpret scientific discovery in a particular way. It has provided valuable insights and these have been used to advance biological science in novel ways. I am not one of those critics of 'the selfish gene' idea who deny its impact and value. But it is nevertheless a metaphor. It is not a straightforward empirical scientific hypothesis. To demonstrate this I want to challenge the reader to a thought experiment. I will first give you one of the central statements of the 'selfish gene' idea. I will then rewrite it so that each sub-phrase (except for one anodyne statement) is replaced by a possible alternative, founded on an opposing metaphor that will form the basis for the rest of this book. The challenge is to think of an empirical test that could possibly distinguish between these two diametrically opposed ways of seeing the relationship between genes and phenotypes.

First, then, the original 'selfish gene' statement (Dawkins 1976: 21):

> Now they swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by tortuous indirect routes, manipulating it by remote control. They are in you and me; they created us, body and mind; and their preservation is the ultimate rationale for our existence.

I would like the reader to think carefully about this statement to absorb its full import. Ask yourself whether you find the statement self-evident, shocking, implausible, likely, true, false, nonsense. Is it theory, fact, or neither? Form a view about it before you continue. Whichever of these views you hold (and all have been expressed by readers of *The selfish gene*), I believe you will find the test an interesting challenge.

So, now let's see what happens when we replace each phrase, except for the phrase 'they are in you and me', by an alternative written from an opposing viewpoint, that of 'genes as prisoners':

> Now they are trapped in huge colonies, locked inside highly intelligent beings, moulded by the outside world, communicating with it by complex processes, through which, blindly, as if by magic, function emerges. They are in you and me; we are the system that allows their code to be read; and their preservation is totally dependent on the joy we experience in reproducing ourselves. We are the ultimate rationale for their existence.

The experiment is made even more effective if we arrange the two statements in register: