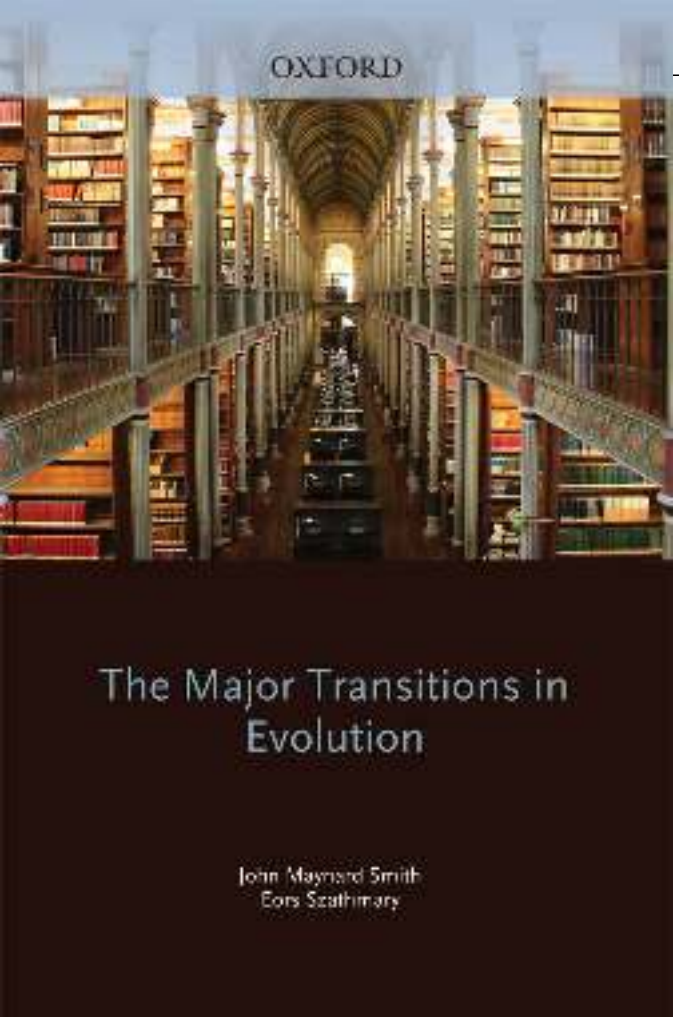
OXFORD
The Major Transitions in Evolution
John Maynard Smith
Eors Szathmary

This page intentionally left blank

THE MAJOR TRANSITIONS IN EVOLUTION

John Maynard Smith and Eörs Szathmáry

OXFORD
UNIVERSITY PRESS

This book has been printed digitally and produced to a standard specification
in order to ensure its continuing availability

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 South Korea Poland Portugal
Singapore 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

Reprinted 2010

ISBN 978-0-19-850294-4

To our parents

This page intentionally left blank

Contents

Tables

This page intentionally left blank

Preface

This book is about the origin of life, of the genetic code, of cells, of sex, of multicellular organisms, of societies, and of language. Such a book is inevitably speculative, because it is an account of a series of unique events that happened a long time ago. But these are matters on which we must speculate. Why else would we study evolution? It is true that evolutionary biology has some practical relevance — for example, to animal breeding or to the origins of antibiotic resistance — but the real reason why we study it is that we are interested in origins. We want to know where we came from.

Although the book is speculative, however, we think that it is a contribution to science, and not to fantasy. Speculation is constrained in two ways. First, each event must be explained in a way that is consistent with a general theory of evolutionary change, the theory of evolution by natural selection. Second, an adequate account of the origin of any system must explain the peculiarities of that system as it exists today: for example, a theory of the origin of the code should explain why it is a triplet code, why it is redundant, why similar codons specify chemically similar amino acids, and so on. In other words, theories about origins can be tested by looking at the present.

This means that we have had to think very hard about the essential characteristics of the various levels of organization whose origins we are trying to explain. We have had to learn a lot of biology, from molecular genetics to linguistics, to write the book. We think that anyone who reads it will also learn some biology.

We have to thank Barry Cox for catalysing our first encounter at the ICSEB III in 1985 at Brighton. The basic ideas behind the book first emerged in 1987–8, when E.S. spent some time with J.M.S. at the University of Sussex. The plan actually to write a book, however, was not conceived until 1991, when E.S. was at the National Institute for Medical Research in London. It seemed that we might be able to collaborate fruitfully because we have a common conviction that heredity and selection are central to evolutionary explanations, but otherwise tend to approach biological problems from different ends — respectively, from chemistry and the physical sciences, and from natural history. Most of the time we have been working in different countries. We have shared the job of writing preliminary drafts of the various chapters, but the final version of every chapter is very much a joint effort, which we have argued about on the many occasions when we have met. There remain a few disagreements between us about what should have been included — E.S. always wanted more

chemistry than J.M.S. could understand — but no significant disagreements about ideas.

While the book was being written, E.S. was a guest at various institutions abroad. Tom Kirkwood, then head of the Laboratory of Mathematical Biology at the National Institute for Medical Research, was very tolerant of this not very MRC-like activity, and he was also a keen discussant. The Institute for Advanced Study in Berlin (Wissenschaftskolleg zu Berlin) provided E.S. with luxurious working and financial conditions for 10 months, where he could discuss many topics with Peter Hammerstein and James Griesemer. Finally, a guest professorship at the Institute of Zoology, University of Zürich, gave the opportunity to teach a course on theoretical biology, largely based on the manuscript of this book. Rüdiger Wehner, as head of the institute, supported this strongly.

The colleagues with whom we have discussed particular problems are too numerous to mention individually, but a few must be named. Before starting on the chapters dealing with development, we had a splendid two-day tutorial with Jonathon Cooke. An earlier draft, excepting the last two chapters, was read by Laurence Hurst, Stephen Kearsey and Mark Ridley. Their comments were invaluable, not only in pointing out errors but in suggesting ways in which things could be explained more clearly. We have not always taken their advice: it is impossible to resist mentioning one example when we did not. We drew a parallel between Eigen's notion of an error threshold and a phase transition. One of the three wrote 'This does NOT remind me of a phase transition: it cannot be too widely known that nothing reminds me of a phase transition'. Behaving, as we do, that the only thing required to make this particular commentator the complete evolutionary biologist is a love of phase transitions, we have left the remark as it stood.

Michael Rodgers has been unfailingly helpful, not least in finding three such wise and knowledgeable referees. He was also responsible for finding Sarah Bunney, who copyedited the manuscript with great care, and on occasions made helpful suggestions on matters of substance, and Jane Templeman, who prepared the figures from our often rather scruffy sketches.

Our greatest debt, however, is to the two men, J. B. S. Haldane and Tibor Gánti, who taught us. Some of their ideas are referred to in the text, but their influence is present on every page. Without their teaching, we might never have tried to write the book at all.

1 Introduction

This page intentionally left blank

1.1 Preamble

Living organisms are highly complex, and are composed of parts that function to ensure the survival and reproduction of the whole. This book is about how and why this complexity has increased in the course of evolution. The increase has been neither universal nor inevitable. Bacteria, for example, are probably no more complex today than their ancestors 2000 million years ago. The most that we can say is that some lineages have become more complex in the course of time. Complexity is hard to define or to measure, but there is surely some sense in which elephants and oak trees are more complex than bacteria, and bacteria than the first replicating molecules.

Our thesis is that the increase has depended on a small number of major transitions in the way in which genetic information is transmitted between generations. Some of these transitions were unique: for example, the origin of the eukaryotes from the prokaryotes, of meiotic sex, and of the genetic code itself. Other transitions, such as the origin of multicellularity, and of animal societies, have occurred several times independently. There is no reason to regard the unique transitions as the inevitable result of some general law: one can imagine that life might have got stuck at the prokaryote or at the protist stage of evolution.

There are obvious difficulties in discussing unique events that happened a long time ago. How can we ever know that our suggested explanations are correct? After all, historians cannot agree about the causes of the Second World War. We accept that certainty is impossible, but there are several reasons why we think the enterprise is worth while. First, we have one great advantage over historians: we have agreed theories both of chemistry and of the mechanism of evolutionary change. We can therefore insist that our explanations be plausible both chemically, and in terms of natural selection. This places a severe constraint on possible theories. Indeed, the difficulty often lies not in choosing between rival theories, but in finding any theory that is chemically and selectively plausible. Further, theories are often testable by looking at existing organisms.

A second reason why the study of origins is worth while is that, to understand the origin of some structure, one must first understand what is essential about it—what features it must have if it is to work at all. Writing this book has forced us to learn a lot of biology, from the nature of the genetic code to the nature of human language. But the major reason for thinking about origins is more intangible: we want to know the answers.

In this introduction, we give an outline of the rest of the book. But, first, we discuss two preliminary topics. In section 1.2, we explain why we do not regard the evolutionary process as one of inevitable progress; in section 1.3, we ask how complexity might be defined and measured. Unfortunately, no very illuminating answers to these questions are possible. Readers who already know that evo-

lution is not equivalent to progress, and that complexity is hard to measure, can safely skip to the start of section 1.3.

In section 1.3, we list the major transitions. The justification for discussing such an apparently diverse series of changes in a single book is that they have features in common. The most important of these is that entities that were capable of independent replication before the transition can replicate only as part of a larger whole after it. This common feature raises a common problem: why does not selection between entities at the lower level disrupt integration at the higher one? These points are explained in more detail in section 1.5. Other features that are common to different transitions are described in section 1.6.

1.2 The fallacy of progress

The notion of progress has a bad name among evolutionary biologists. Lamarck accepted the earlier idea of a ladder of nature, and argued that organisms have an inherent tendency to climb the ladder. It was Lamarck's notion of an inherent tendency, rather than his belief in the inheritance of acquired characters, that Darwin was rejecting when he said that his theory had nothing in common with Lamarck's: he rightly saw that to explain evolution by an inherent tendency is as vacuous as to say that a man is fat because he has an inherent tendency to obesity. Today, we are unhappy with a picture of evolution that places us at the summit, and arranges all other organisms in a line behind us: what have we to be so proud about? To be fair, humans were by no means at the summit of the medieval scala naturae; there were angels and archangels above us as well as worms below.

There are, of course, more solid reasons, both empirical and theoretical, for rejecting a simple image of progress on a linear scale. Empirically, the history of life is better visualized as a branching tree than as a single ascending line. The fossil record shows that many organisms – horseshoe crabs, the coelacanth, crocodiles, for example – have undergone little change, progressive or otherwise, for hundreds of millions of years. On a shorter timescale, sibling species tell the same story. The fruit flies *Drosophila melanogaster* and *D. simulans* are hard to distinguish morphologically, but molecular data indicate that they are separated by several million years of evolution. Hence, either morphological evolution in the two species has been almost exactly parallel, which is implausible, or neither species has changed.

On the theoretical side, there is no reason why evolution by natural selection should lead to an increase in complexity, if that is what we mean by progress. At most, the theory suggests that organisms should get better, or at least no worse, at doing what they are doing right now. But an increase in 'maximum fitness' – that is, expected number of offspring – may be achieved by losing eyes or legs as well as by gaining them. Even if an increase in fitness cannot be equated with

an increase in complexity, or with progress, it might seem at first sight that R. A. Fisher's (1930) 'fundamental theorem of natural selection' at least guarantees an increase in fitness. The theorem states that the rate of increase in the overall fitness of a population is equal to the genetic variance in fitness: since variances cannot be negative, the theorem states that fitness can only increase. If so, 'mean fitness' in biology is an analogue of entropy in physics: it gives an arrow to time. Thus in physics the inevitable increase in entropy distinguishes past from future: if mean fitness can only increase, this gives a direction to evolution. It seems that Fisher did indeed think that his theorem could play such a role: otherwise, why 'fundamental'? Unfortunately, the theorem holds only if the relative fitnesses of genotypes are constant, and independent of their frequencies in the population: for many traits, such constancy does not hold.

1.3 The measurement of complexity

Even if progress is not a universal law of evolution, common sense does suggest that at least some lineages have become more complex. How might one measure this increase? A possible answer is in terms of the DNA content of the genome, which can be thought of as instructions for making the organism: more complex organisms require lengthier instructions. If we look at total DNA, we reach the rather demeaning conclusion that a lungfish or a lily is some 40 times as complex as a human (Table 1.1). Things make more sense if we allow for the fact that a surprisingly small proportion of the DNA codes for anything. It then appears that eukaryotes have more coding DNA than prokaryotes (although the difference between yeast and *Escherichia coli* is small); that higher plants and invertebrate animals have more DNA than single-celled organisms; that an invertebrate with wings, legs and eyes has more DNA than a nematode; and that

Table 1.1 Genome size and DNA content

	Genome size (base pairs $\times 10^9$)	Coding DNA (%)
Bacterium (*Escherichia coli*)	0.004	100
Yeast (*Saccharomyces*)	0.009	70
Nematode (*Caenorhabditis*)	0.09	25
Fruit fly (*Drosophila*)	0.18	33
Newt (*Triturus*)	19.0	1.5–4.5
Human	3.5	9–27
Lungfish (*Protopterus*)	140.0	0.4–1.2
Flowering plant (*Arabidopsis*)	0.2	31
Flowering plant (*Fritillaria*)	130.0	0.02

Data from Cavalier-Smith, 1985.

vertebrates have more DNA than invertebrates. This last observation is rather puzzling: it is possible that the large brains of vertebrates require large genomes.

These data make sense, but what they tell us about structural and functional complexity is very limited. At present, we can say rather little about how much genetic information is needed to program a particular morphological structure.

1.4 The major transitions

A tentative list of the major stages in the evolution of complexity, and the transitions between them, is given in Table 1.2. We have confined our attention to the way in which information is transmitted between generations, so have not included major phenotypic changes, such as the conquest of land by plants and animals, or the origins of vision, or of flight or of homiothermy, which did not involve such a change in the method of information transmission.

One feature is common to many of the transitions: entities that were capable of independent replication before the transition can replicate only as part of a larger whole after it. Some examples will make this clearer:

The origin of chromosomes. Initially, there were independently replicating nucleic acid molecules: after the transition, a set of linked molecules must replicate together.

The origin of eukaryotes. The ancestors of mitochondria and chloroplasts were once free-living prokaryotes: today, they can replicate only within a host cell.

The origin of sex. The first eukaryotes could, presumably, reproduce asexually, and on their own: today, most eukaryotes can replicate only as part of a sexual population.

The origin of multicellular organisms. The cells of animals, plants and fungi are descended from single-celled protists, each of which could survive on its own: today, they exist (outside the laboratory) only as parts of larger organisms.

Table 1.2 The major transitions

Replicating molecules	—	Populations of molecules in compartments
Independent replicators	→	Chromosomes
RNA as gene and enzyme	→	DNA + protein (genetic code)
Prokaryotes	→	Eukaryotes
Asexual clones	→	Sexual populations
Protists	→	Animals, plants, fungi (cell differentiation)
Solitary individuals	→	Colonies (non-representative cells)
Primate societies	→	Human society (language)

The origin of social groups. Individual ants, bees, wasps and termites can survive and transmit genes (either their own or ones genetically similar to their own) only as part of a social group: the same is effectively true of humans.

It might be asked why we do not include the origin of ecosystems in our list of transitions. There are two reasons. First, in temporal terms an ecosystem is not the final stage in a series: ecosystems are as old as replicating molecules. Second, ecosystems are not individuals, separated from others, whereas the other stages we have listed (including sexual species, and insect colonies) do have a degree of individuality, and separateness from other entities of the same kind. For this reason, ecosystems cannot be units of selection.

Given this common feature of the major transitions, there is a common question we can ask of them. Why did not natural selection, acting on entities at the lower level (replicating molecules, free-living prokaryotes, asexual protists, single cells, individual organisms), disrupt integration at the higher level (chromosomes, eukaryotic cells, sexual species, multicellular organisms, societies)? It is because there is this common question that we have found it illuminating to compare the different transitions. In fact, one of the stimuli for attempting the work was our realization that a model one of us had developed to analyse the origin of compartments containing populations of molecules was formally and mathematically similar to a model that the other had developed to analyse the evolution of cooperative behaviour in higher animals.

First, we must establish that the problem is not an imaginary one: that there is a real danger that selection at the lower level will disrupt integration at the higher. This is best done by listing examples (Table 1.3) in which such a process can be observed today:

- If Mendel's laws are rigorously obeyed, a gene can only increase its representation in future generations by ensuring the success of the cell in which it finds itself, and of the other genes in the cell. Hence, Mendel's laws ensure the evolution of cooperative, or 'coadapted', genes. But the laws are broken, by meiotic drive, and by transposable elements. These are examples of the more general phenomenon of intragenomic conflict, which is the topic of Chapter 10.

Table 1.3 Conflict between selection at different levels

Form of cooperation	Exceptions
Fair meiosis	Meiotic drive, transposition
Sexual reproduction	Parthenogenesis
Differentiation of somatic cells	Escape from growth control
Non-reproductive castes of social insects	Egg-laying worker bees

- A sexual population has an advantage, in rate of evolution and in the elimination of harmful mutations, over an asexual one. But a parthenogenetic female has, in the short run, a twofold advantage over a sexual one, and parthenogens are not uncommon.
- A gene in a somatic cell of a plant might best ensure the transmission of replicas of itself by giving rise to a flower bud, even if this reduced the success of the whole plant.
- A bee colony produces more reproductives if the workers raise the queen's offspring. But workers do lay eggs (which are unfertilized, and hence male).

We cannot hope to explain these transitions in terms of the ultimate benefits they conferred. For example, it may be that, in the long run, the most important difference between prokaryotes and eukaryotes is that the latter evolved a mechanism for chromosome segregation at cell division that permits DNA replication to start simultaneously at many origins, whereas prokaryotes have only a single origin of replication. At the very least, this was a necessary precondition for the subsequent increase in DNA content, without which complexity could not increase. But this is not the reason why the change occurred in the first place; as we explain in Chapter 6, the new segregation mechanism was forced on the early eukaryotes by the loss of a rigid cell wall, which plays a crucial role in the segregation of prokaryotic chromosomes. Or to take a second example, meiotic sex was an important preadaptation for the subsequent evolutionary radiation of the eukaryotes, but it could not have originated for that reason.

The transitions must be explained in terms of immediate selective advantage to individual replicators: we are committed to the gene-centred approach outlined by Williams (1966), and made still more explicit by Dawkins (1976). There is, in fact, one feature of the transitions listed in Table 1.1 that leads to this conclusion. At some point in the life cycle, there is only one copy, or very few copies, of the genetic material: consequently, there is a high degree of genetic relatedness between the units that combine in the higher organism. The importance of this general principle was first emphasized by Hamilton (1964) in his explanation of the evolution of social behaviour, but we believe it to be quite general. To give two other examples: multicellular organisms develop from a single fertilized egg, so that their cells are genetically identical, except for somatic mutations; most eukaryotes inherit their organelles from one parent only, so that the organelles in a single individual are almost always genetically identical. We think that a similar principle operated in the origin of the earliest cells.

The principle of genetic similarity resulting from a small number of founders is important at the time of the transition. Two other processes—contingent irreversibility and central control—help to explain the maintenance of higher-level entities, once they have arisen, although they are less relevant to the origin of such entities.

Contingent irreversibility

If an entity has replicated as part of a larger whole for a long time, it may have lost the capacity for independent replication that it once had, for accidental reasons that have little to do with the selective forces that led to the evolution of the higher level entity in the first place. For example, mitochondria cannot resume independent existence, if only because most of their genes have been transferred to the nucleus; cancer cells may escape growth control, but have no independent future as protists; worker bees may lay male eggs, but cannot establish a new colony on their own.

The contingent nature of irreversibility is perhaps best illustrated by the reversion from sex to parthenogenesis. Mammals are never parthenogens, probably because, at some but in some tissues, only the allele inherited from the father is active; hence, in an embryo with no father, some essential gene activities are missing. Gymnosperms are also never parthenogens, perhaps for a different reason: chloroplasts are transmitted in the pollen. Anamniote vertebrates (fish and amphibians), although they may be parthenogens, always require sperm from the males of another species to initiate development, perhaps because the sperm provides a centriole. The relevance of these sexual hangups — and there are many others — is to show how various and accidental are the reasons why reversal is difficult or impossible.

Three points must be made about irreversibility:

- Different mechanisms may be involved in the origin and the maintenance of higher level organization. It would be absurd to suggest sexual imprinting as a cause of the origin of sex, or gene transfer to the nucleus as a cause of the symbiotic origin of mitochondria.
- Irreversibility is not absolute. There are, after all, many successful sexual parthenogens. To give a second example, viruses and transposons are probably the descendents of well-behaved chromosomal genes. Irreversibility, therefore, is not only irrelevant to the origin of higher-level organization: it is not a sufficient explanation for its maintenance.
- The fact that a once independent replicator can no longer revert does not mean that selfish mutations cannot occur in the genes of such replicators.

Central control

Analogy with human society might suggest that the maintenance of organization depends on some kind of central control: people pay taxes because they would be punished if they did not. But the analogy is not a close one. In human societies, central control depends either on the existence of an armed group within society — for example, a feudal aristocracy — or on a consensus imposed by a majority, or on some combination of the two. The notion of armed force has little application in biology — there are not armed and unarmed genes. Although