NICK LANE

THE VITAL QUESTION



WHY IS LIFE THE WAY IT IS?

PROFILE BOOKS

First published in Great Britain in 2015 by
PROFILE BOOKS LTD
3 Holford Yard
Bevin Way
London
WCIX 9HD
www.profilebooks.com

Copyright © Nick Lane, 2015

1 3 5 7 9 10 8 6 4 2

Printed and bound in Great Britain by Clays, Bungay, Suffolk

Typeset in Fournier by MacGuru Ltd info@macguru.org.uk

The moral right of the author has been asserted.

All rights reserved. Without limiting the rights under copyright reserved above, no part of this publication may be reproduced, stored or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise), without the prior written permission of both the copyright owner and the publisher of this book.

 \boldsymbol{A} CIP catalogue record for this book is available from the British Library.

ISBN 978 1 78125 036 5 eISBN 978 1 84765 880 7



CONTENTS

	Introduction: Why is life the way	it is? 1
	Part I: The Problem	
I	What is life?	19
2	What is living?	53
	Part II: The Origin of Life	
3	Energy at life's origin	89
4	The emergence of cells	122
	Part III: Complexity	
5	The origin of complex cells	157
6	Sex and the origins of death	192
	Part IV: Predictions	
7	The power and the glory	237
	Epilogue: From the deep	281
	Glossary	291
	Acknowledgements	300
	Bibliography	306
	List of illustrations	334
	Index	338

INTRODUCTION: WHY IS LIFE THE WAY IT IS?



There is a black hole at the heart of biology. Bluntly put, we do not f L know why life is the way it is. All complex life on earth shares a common ancestor, a cell that arose from simple bacterial progenitors on just one occasion in 4 billion years. Was this a freak accident, or did other 'experiments' in the evolution of complexity fail? We don't know. We do know that this common ancestor was already a very complex cell. It had more or less the same sophistication as one of your cells, and it passed this great complexity on not just to you and me but to all its descendants, from trees to bees. I challenge you to look at one of your own cells down a microscope and distinguish it from the cells of a mushroom. They are practically identical. I don't live much like a mushroom, so why are my cells so similar? It's not just that they look alike. All complex life shares an astonishing catalogue of elaborate traits, from sex to cell suicide to senescence, none of which is seen in a comparable form in bacteria. There is no agreement about why so many unique traits accumulated in that single ancestor, or why none of them shows any sign of evolving independently in bacteria. Why, if all of these traits arose by natural selection, in which each step offers some small advantage, did equivalent traits not arise on other occasions in various bacterial groups?

These questions highlight the peculiar evolutionary trajectory of life on earth. Life arose around half a billion years after the earth's formation,

perhaps 4 billion years ago, but then got stuck at the bacterial level of complexity for more than 2 billion years, half the age of our planet. Indeed, bacteria have remained simple in their morphology (but not their biochemistry) throughout 4 billion years. In stark contrast, all morphologically complex organisms – all plants, animals, fungi, seaweeds and single-celled 'protists' such as amoeba – descend from that singular ancestor about 1.5–2 billion years ago. This ancestor was recognisably a 'modern' cell, with an exquisite internal structure and unprecedented molecular dynamism, all driven by sophisticated nanomachines encoded by thousands of new genes that are largely unknown in bacteria. There are no surviving evolutionary intermediates, no 'missing links' to give any indication of how or why these complex traits arose, just an unexplained void between the morphological simplicity of bacteria and the awesome complexity of everything else. An evolutionary black hole.

We spend billions of dollars a year on biomedical research, teasing out the answers to unimaginably complex questions about why we get ill. We know in enormous detail how genes and proteins relate to each other, how regulatory networks feed back into each other. We build elaborate mathematical models and design computer simulations to play out our projections. Yet we don't know how the parts evolved! How can we hope to understand disease if we have no idea why cells work the way they do? We can't understand society if we know nothing of its history; nor can we understand the workings of the cell if we don't know how it evolved. This isn't just a matter of practical importance. These are human questions about why we are here. What laws gave rise to the universe, the stars, the sun, the earth, and life itself? Will the same laws beget life elsewhere in the universe? Would alien life be anything like us? Such metaphysical questions lie at the heart of what makes us human. Some 350 years after the discovery of cells, we still don't know why life on earth is the way it is.

You might not have noticed that we don't know. It's not your fault. Text books and journals are full of information, but often fail to address these 'childlike' questions. The internet swamps us with all manner of indiscriminate facts, mixed with varying proportions of nonsense. But it's not merely a case of information overload. Few biologists are more than dimly aware

of the black hole at the heart of their subject. Most work on other questions. The great majority study large organisms, particular groups of plants or animals. Relatively few work on microbes, and even fewer on the early evolution of cells. There's also a concern about creationists and intelligent design — to admit we don't know all the answers risks opening the door to naysayers, who deny that we have any meaningful knowledge of evolution. Of course we do. We know an awful lot. Hypotheses on the origins of life and the early evolution of cells must explain an encyclopedia of facts, conform to a straitjacket of knowledge, as well as predict unexpected relationships that can be tested empirically. We understand a great deal about natural selection and some of the more random processes that sculpt genomes. All these facts are consistent with the evolution of cells. But this same straitjacket of facts is precisely what raises the problem. We don't know why life took the peculiar course that it did.

Scientists are curious people, and if this problem were as stark as I'm suggesting, it would be well known. The fact is – it's far from obvious. The various competing answers are esoteric, and all but obscure the question. Then there's the problem that clues come from many disparate disciplines, from biochemistry, geology, phylogenetics, ecology, chemistry and cosmology. Few can claim real expertise in all those fields. And now we are in the midst of a genomic revolution. We have thousands of complete genome sequences, codes that stretch over millions or billions of digits, all too often containing conflicting signals from the deep past. Interpreting these data demands rigorous logical, computational and statistical know-how; any biological understanding is a bonus. And so the clouds have been swirling around with arguments. Each time a gap opens up, it reveals an increasingly surreal landscape. The old comforts have been evaporating. We're now faced with a stark new picture, and it's both real and troubling. And from a researcher's standpoint, hoping to find some significant new problem to solve, it's flat out thrilling! The biggest questions in biology are yet to be solved. This book is my own attempt to make a start.

How do bacteria relate to complex life? The roots of the question date right back to the discovery of microbes by the Dutch microscopist Antony van Leeuwenhoek in the 1670s. His menagerie of 'little animals' thriving

under the microscope took some believing, but was soon confirmed by the equally ingenious Robert Hooke. Leeuwenhoek also discovered bacteria, and wrote about them in a famous paper of 1677: they were 'incredibly small; nay, so small, in my sight, that I judged that even if 100 of these very wee animals lay stretched out one against another, they could not reach the length of a grain of course sand; and if this be true, then ten hundred thousand of these living creatures could scarce equal the bulk of a course grain of sand'. Many doubted that Leeuwenhoek could possibly have seen bacteria using his simple single-lens microscopes, though it is now incontrovertible that he did so. Two points stand out. He found bacteria everywhere – in rainwater and the sea, not just on his own teeth. And he intuitively made some distinction between these 'very wee animals' and the 'gygantick monsters' - microscopic protists! - with their enthralling behaviour and 'little feet' (cilia). He even noticed that some larger cells were composed of a number of little 'globules', which he compared with bacteria (though not in those terms). Among these little globules, Leeuwenhoek almost certainly saw the cell nucleus, the repository of the genes in all complex cells. And there the matter lay for several centuries. The famous classifier Carl Linnaeus, 50 years after Leeuwenhoek's discoveries, had just lumped all microbes into the genus *Chaos* (formless) of the phylum Vermes (worms). In the nineteenth century Ernst Haeckel, the great German evolutionist and contemporary of Darwin, formalised the deep distinction again, separating bacteria from other microbes. But in conceptual terms, there was little advance until the middle of the twentieth century.

The unification of biochemistry brought matters to a head. The sheer metabolic virtuosity of bacteria had made them seem uncategorisable. They can grow on anything, from concrete to battery acid to gases. If these totally different ways of life had nothing in common, how could bacteria be classified? And if not classified, how could we understand them? Just as the periodic table brought coherence to chemistry, so biochemistry brought an order to the evolution of cells. Another Dutchman, Albert Kluyver, showed that similar biochemical processes underpinned the extraordinary diversity of life. Processes as distinct as respiration, fermentation and photosynthesis all shared a common basis, a conceptual integrity which attested that all life

had descended from a common ancestor. What was true of bacteria, he said, was also true of elephants. At the level of their biochemistry, the barrier between bacteria and complex cells barely exists. Bacteria are enormously more versatile, but the basic processes that keep them alive are similar. Kluyver's own student Cornelis van Niel, together with Roger Stanier, perhaps came closest to appreciating the difference: bacteria, like atoms, could not be broken down any further, they said: bacteria are the smallest unit of function. Many bacteria can respire oxygen in the same way that we do, for example, but it takes the whole bacterium to do so. Unlike our own cells, there are no internal parts dedicated to respiration. Bacteria divide in half as they grow, but in function they are indivisible.

And then came the first of three major revolutions that have wracked our view of life in the past half century. This first was instigated by Lynn Margulis in the summer of love, 1967. Complex cells did not evolve by 'standard' natural selection, Margulis argued, but in an orgy of cooperation, in which cells engaged with each other so closely that they even got inside each other. Symbiosis is a long-term interaction between two or more species, usually some sort of trade for wares or services. In the case of microbes, those wares are the substances of life, the substrates of metabolism, which power the lives of cells. Margulis talked about *endo*symbiosis – the same types of trade, but now so intimate that some collaborating cells physically live inside their host cell like the traders who sold from within the temple. These ideas trace their roots to the early twentieth century, and are reminiscent of plate tectonics. It 'looks' as if Africa and South America were once joined together, and later pulled apart, but this childlike notion was long ridiculed as absurd. Likewise, some of the structures inside complex cells look like bacteria, and even give the impression of growing and dividing independently. Perhaps the explanation really was as simple as that - they are bacteria!

Like tectonics these ideas were ahead of their time, and it was not until the era of molecular biology in the 1960s that it was possible to present a strong case. This Margulis did for two specialised structures inside cells — the mitochondria, seats of respiration, in which food is burned in oxygen to provide the energy needed for living, and the chloroplasts, the engines of

photosynthesis in plants, which convert solar power into chemical energy. Both of these 'organelles' (literally miniature organs) retain tiny specialised genomes of their own, each one with a handful of genes encoding at most a few dozen proteins involved in the mechanics of respiration or photosynthesis. The exact sequences of these genes ultimately gave the game away—plainly, mitochondria and chloroplasts do derive from bacteria. But notice I say 'derive'. They are no longer bacteria, and don't have any real independence, as the vast majority of the genes needed for their existence (at least 1,500 of them) are found in the nucleus, the genetic 'control centre' of the cell.

Margulis was right about the mitochondria and chloroplasts; by the 1980s, few doubters remained. But her enterprise was much greater: for Margulis, the entire complex cell, now generally known as the eukaryotic cell (from the Greek meaning 'true nucleus') was a patchwork of symbioses. In her eyes, many other parts of the complex cell, notably the cilia (Leeuwenhoek's 'little feet'), also derived from bacteria (spirochetes in the case of cilia). There had been a long succession of mergers, which Margulis now formalised as the 'serial endosymbiosis theory'. Not just individual cells but the whole world was a vast collaborative network of bacteria – 'Gaia', an idea that she pioneered with James Lovelock. While the concept of Gaia has enjoyed a renaissance in the more formal guise of 'earth systems science' in recent years (stripping Lovelock's original teleology), the idea that complex 'eukaryotic' cells are an ensemble of bacteria has far less to support it. Most of the structures of the cell do not look as if they derived from bacteria, and there's nothing in the genes to suggest that they do. So Margulis was right about some things and almost certainly wrong about others. But her crusading spirit, forceful femininity, dismissal of Darwinian competition and tendency to believe in conspiracy theories, meant that when she died prematurely from a stroke in 2011, her legacy was decidedly mixed. A feminist heroine for some and loose cannon for others, much of this legacy was sadly far removed from science.

Revolution number two was the phylogenetic revolution – the ancestry of genes. The possibility had been anticipated by Francis Crick as early as 1958. With characteristic aplomb, he wrote: 'Biologists should realise that

before long we shall have a subject which might be called "protein taxonomy" – the study of amino acid sequences of proteins of an organism and the comparison of them between species. It can be argued that these sequences are the most delicate expression possible of the phenotype of an organism and that vast amounts of evolutionary information may be hidden away within them.' And lo, it came to pass. Biology is now very much about the information concealed in the sequences of proteins and genes. We no longer compare the sequences of amino acids directly, but the sequences of letters in DNA (which encodes proteins), giving even greater sensitivity. Yet for all his vision, neither Crick nor anyone else began to imagine the secrets that actually emerged from the genes.

The scarred revolutionary was Carl Woese. In work beginning quietly in the 1960s and not bearing fruit until a decade later, Woese selected a single gene to compare between species. Obviously, the gene had to be present in all species. What's more, it had to serve the same purpose. That purpose had to be so fundamental, so important to the cell, that even slight changes in its function would be penalised by natural selection. If most changes are eliminated, what remains must be relatively unchanging — evolving extremely slowly, and changing little over vast periods of time. That's necessary if we want to compare the differences that accumulate between species over literally billions of years, to build a great tree of life, going back to the beginning. That was the scale of Woese's ambition. Bearing all these requirements in mind, he turned to a basic property of all cells, the ability to make proteins.

Proteins are assembled on remarkable nanomachines found in all cells, called ribosomes. Excepting the iconic double helix of DNA, nothing is more symbolic of the informational age of biology than the ribosome. Its structure also epitomises a contradiction that is hard for the human mind to fathom—scale. The ribosome is unimaginably tiny. Cells are already microscopic. We had no inkling of their existence for most of human history. Ribosomes are orders of magnitude smaller still. You have 13 *million* of them in a single cell from your liver. But ribosomes are not only incomprehensibly small; on the scale of atoms, they are massive, sophisticated superstructures. They're composed of scores of substantial subunits, moving

machine parts that act with far more precision than an automated factory line. That's not an exaggeration. They draw in the 'tickertape' code-script that encodes a protein, and translate its sequence precisely, letter by letter, into the protein itself. To do so, they recruit all the building blocks (amino acids) needed, and link them together into a long chain, their order specified by the code-script. Ribosomes have an error rate of about one letter in 10,000, far lower than the defect rate in our own high-quality manufacturing processes. And they operate at a rate of about 10 amino acids per second, building whole proteins with chains comprising hundreds of amino acids in less than a minute. Woese chose one of the subunits from the ribosome, a single machine part, so to speak, and compared its sequence across different species, from bacteria such as *E. coli* to yeast to humans.

His findings were a revelation, and turned our world view on its head. He could distinguish between the bacteria and complex eukaryotes without any difficulty, laying out the branching tree of genetic relatedness within and between each of these magisterial groups. The only surprise in this was how little difference there is between plants and animals and fungi, the groups that most biologists have spent most of their lives studying. What nobody anticipated was the existence of a third domain of life. Some of these simple cells had been known for centuries, but were mistaken for bacteria. They look like bacteria. Exactly like bacteria: equally tiny, and equally lacking in discernible structure. But the difference in their ribosomes was like the smile of the Cheshire cat, betraying the presence of a different sort of absence. This new group might have lacked the complexity of eukaryotes, but the genes and proteins that they did have were shockingly different from those of bacteria. This second group of simple cells became known as the archaea, on the hunch that they're even older than the bacteria, which is probably not true; modern views have it that they are equally old. But at the arcane level of their genes and biochemistry, the gulf between bacteria and archaea is as great as that between bacteria and eukaryotes (us). Almost literally. In Woese's famous 'three domains' tree of life, archaea and eukaryotes are 'sister groups', sharing a relatively recent ancestor.

In some respects, the archaea and eukaryotes do indeed have a lot in common, especially in terms of information flow (the way that they read off their genes and convert them into proteins). In essence, archaea have a few sophisticated molecular machines resembling those of eukaryotes, if with fewer parts – the seeds of eukaryotic complexity. Woese refused to countenance any deep morphological gulf between bacteria and eukaryotes, but proposed three equivalent domains, each of which had explored vast realms of evolutionary space, none of which could be given precedence. Most forcefully, he rejected the old term 'prokaryote' (meaning literally 'before the nucleus', which could be applied to both archaea and bacteria) as there was nothing in his tree to suggest a genetic basis for that distinction. On the contrary, he pictured all three domains as reaching right back into the deepest past, sharing a mysterious common ancestor, from which they had somehow 'crystallised'. Towards the end of his life Woese became almost mystical about those earliest stages of evolution, calling for a more holistic view of life. That's ironic, given that the revolution he wrought was based on a wholly reductionist analysis of a single gene. There is no doubt that the bacteria, archaea and eukaryotes are genuinely distinct groups and that Woese's revolution was real; but his prescription for holism, taking whole organisms and full genomes into account, is right now ushering in the third cellular revolution – and it overturns Woese's own.

This third revolution isn't over yet. It's a little more subtle in reasoning, but packs the biggest punch of all. It is rooted in the first two revolutions, and specifically in the question: how do the two relate? Woese's tree depicts the divergence of one fundamental gene in the three domains of life. Margulis, in contrast, has genes from different species converging together in the mergers and acquisitions of endosymbiosis. Depicted as a tree, this is the fusion, not the bifurcation, of branches – the opposite of Woese. They can't both be right! Neither do they both have to be totally wrong. The truth, as so often in science, lies somewhere between the two. But don't think that makes it a compromise. The answer that's emerging is more exciting than either alternative.

We know that mitochondria and chloroplasts were indeed derived from bacteria by endosymbiosis, but that the other parts of complex cells probably evolved by conventional means. The question is: when, exactly? Chloroplasts are found only in algae and plants, hence were most likely acquired in

an ancestor of those groups alone. That puts them as a relatively late acquisition. Mitochondria, in contrast, are found in all eukaryotes (there's a backstory there that we'll examine in Chapter 1) and so must have been an earlier acquisition. But how early? Put another way, what kind of cell picked up mitochondria? The standard textbook view is that it was quite a sophisticated cell, something like an amoeba, a predator that could crawl around, change shape and engulf other cells by a process known as phagocytosis. In other words, mitochondria were acquired by a cell that was not so far from being a fully fledged, card-carrying eukaryote. We now know that's wrong. Over the last few years, comparisons of large numbers of genes in more representative samples of species have come to the unequivocal conclusion that the host cell was in fact an archaeon – a cell from the domain Archaea. All archaea are prokaryotes. By definition, they don't have a nucleus or sex or any of the other traits of complex life, including phagocytosis. In terms of its morphological complexity, the host cell must have had next to nothing. Then, somehow, it acquired the bacteria that went on to become mitochondria. Only then did it evolve all those complex traits. If so, the singular origin of complex life might have depended on the acquisition of mitochondria. They somehow triggered it.

This radical proposition — complex life arose from a singular endosymbiosis between an archaeon host cell and the bacteria that became mitochondria—was predicted by the brilliantly intuitive and free-thinking evolutionary biologist Bill Martin, in 1998, on the basis of the extraordinary mosaic of genes in eukaryotic cells, a mosaic largely uncovered by Martin himself. Take a single biochemical pathway, say fermentation. Archaea do it one way, and bacteria quite a different way; the genes involved are distinct. Eukaryotes have taken a few genes from bacteria, and a few others from archaea, and woven them together into a tightly knit composite pathway. This intricate fusion of genes doesn't merely apply to fermentation, but to almost all biochemical processes in complex cells. It is an outrageous state of affairs!

Martin thought all this through in great detail. Why did the host cell pick up so many genes from its own endosymbionts, and why did it integrate them so tightly into its own fabric, replacing many of its existing genes in the process? His answer, with Miklós Müller, is called the hydrogen