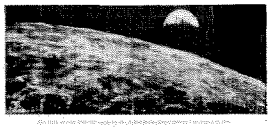
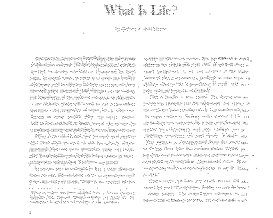
What Is Life? A Closer Look

We can actually classify and enumerate the components of the machine, gene by gene, and discern their interrelated functional organization.

by Robert L. Sinsheimer

(On Seminar Day in April 1967, Robert L. Sinsheimer, then professor of biophysics at Caltech, delivered a talk entitled "What Is Life?" (published in the Caltech Quarterly, *a short-lived spin-off of* E&S*). At the dawn of the era of space exploration, many people were wondering how life was to be recognized if indeed it were found elsewhere than on Earth. Sinsheimer gave the biologist's definition. Now, 30 years later, he sticks by his answer, and fills in some gaps with data from current biological research.)*





The black box (or, more accurately, the colorless capsule) that we call a living cell is a specialized, intricate, highly evolved machine. Nearly three decades ago we were first able to describe qualitatively the essential elements of this self-perpetuating machine (*Caltech Quarterly*, Summer 1967) at the level of its molecular organization. At that time our understanding of the machinery of life was newly emergent, still quite incomplete and porous, but sufficient to replace the older vague theories and speculations.

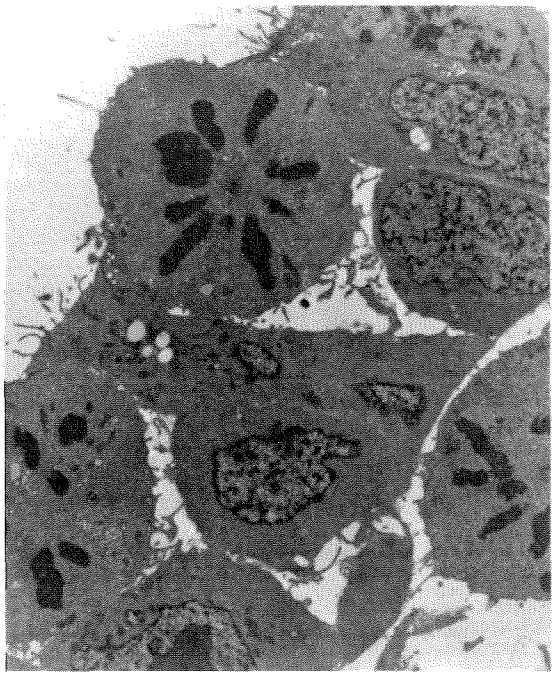
Back in 1967 we were asking: What is the essence of this quality, "life"? Biology is the science of life, but no biology textbooks could provide a definition of what life is. There were two explanations advanced to account for the properties of living beings: one postulated that the substance of living matter was intrinsically different from that of nonliving matter; the other, that the properties of life were solely a consequence of an unusual organization of ordinary matter in living creatures. The latter argument had been hampered by the inability to describe that organization, but by 1967 biologists were able to describe the complex organization of the cell, the basic unit of life, in physical and chemical terms. I wrote then that "what distinguishes life is the presence of a persistent degree of structural complexity at a molecular level—unknown outside the sphere of life." Although higher organisms are composed of millions or billions of cells organized into a cooperative, interacting whole, a single cell can perform all the essential living functions.

Biology has progressed. Within the past year, biologists have anlyzed the complete hereditary instructions—the complete DNA sequences—of two species of microorganisms, *Haemophilus influenzae* Rd and *Mycoplasma genitalium*. Using these two sequenced species (and more are on the way), we can now describe the cell in much more detail. We can actually classify and enumerate the components of the machine, gene by gene, and discern their interrelated functional organization.

But first let me return to sum up the essential features of this specialized molecular organization that I described in 1967:

1. A flexible, self-made bounding membrane, which can be replaced if it is torn and can be enlarged as the cell grows, which defines the ordered, integrated space of life and controls ingress from and egress to the external world. It also provides a two-dimensional surface on which agents involved in particular reaction sequences can arrange themselves.

2. Coordinated groups of specific catalysts (usually proteins) which, by accelerating a myriad



One of the essential features of a living cell is its ability to produce new cells by division and to convey hereditary instructions. This micrograph shows cells in a mouse embryo at 8.5 days of gestation, a time when rapid cell division is taking place. The three round cells with a rosette-like pattern (two of them cut off on the sides) are in different stages of mitosis (cell division). In less than an hour each will have divided into two daughter cells. The dark, sausage-shaped bodies arranged in the "rosette" are the chromosomes into which each cell's DNA is packaged for equal distribution to the daughter cells. Apprixmate magnification: 3000x. (Jean-Paul Revel, Stefan Offermanns, Mel Simon)

In 1967 this qualitative functional description of the elements essential to the organization of a living cell was the best we could do. But now we have, as illustrations, these two completely sequenced microorganisms.

of specific reactions in specific directions, determine the pathways of biosynthesis and biodegradation within the cell.

3. Molecular machinery to convert an external source of energy (either solar energy or the energy strored in the chemical bonds of nutrients) into forms suitable to drive reactions within the cells.

4. A hereditary system of information storage (DNA), which provides in each generation the coded formulas for the structures of the various catalysts and structural members of the cell; also the means to repair the DNA structures if damaged, utilizing the redundancy of information in the double helix; and molecular machinery to convert (by transcription and translation) the inherited coded information in DNA into the specific molecules and structures of the cell. The cell must also have the capability to produce new cells by division, so as to multiply and to permit the trial of modifications of the hereditary instructions.

5. Such division must accurately provide each daughter cell with a full set of the hereditary instructions (occasionally modified by mutation), together with a sufficient endowment of machinery to transcribe and translate these instructions. It must also provide a supply of usable energy adequate to permit each daughter cell to flourish.

6. Because this organization must be flexible and adaptive to changing environments, it has interlocking control systems that automatically regulate the varied functions to keep the cell on course or, more accurately, on its many courses. In addition, the living cell has a capability beyond and distinct from that of the conventional cybernetic system, which can only adjust the rates of action of its component parts to maintain a level of function or output. Since a cell makes all of its component parts, it can also adapt the number of component parts to the task—creating more if needed—and it can also salvage and repair damaged parts if the damage is not too great.

In 1967 this qualitative functional description of the elements essential to the organization of a living cell was the best we could do. But now we have, as illustrations, these two completely sequenced microorganisms.

Haemophilus influenzae Rd is a small, nonmotile bacterium, a pathogen that causes respiratory infections in humans; one strain of it causes meningitis. *H. influenzae* is able to exist in a relatively simple medium and has a genome of 1,830,177 nucleotide pairs (DNA's basic molecular units, whose signifying letters, A, C, G, and T—for adenine, cytosine, guanine, and thymine—make up the genetic code), containing 1,743 regions that are equivalent to genes and that code for specific protein structures. (This work was published by Robert D. Fleischman et al. in *Science*, vol. 269, No. 5223, pp. 496–512, 1995.)

The second species, *Mycoplasma genitalium,* is a tiny, parasitic bacterium, lacking a rigid cell wall and found in association with ciliated epithelial cells of primate and human genital and respiratory tracts. *M. genitalium* has a "stripped-down" genome of 580,070 nucleotide pairs containing only 470 protein-coding regions—likely a near-minimal content for a self-replicating organism, albeit one dependent upon a host for many

complex nutrients and environmental stability (published by Claire M. Fraser et al. in *Science*, vol. 270, No. 5235, pp. 397–403, 1995).

Most, although not all, of the defined gene sequences can be related to proteins of known or likely function, and I thought it would make an interesting exercise to try to fit them into the categories I had described in 1967. The result is shown at left. This doesn't cover the complete organisms: in *Haemophilus*, probable functions have been ascribed to 1,007 of the 1,743; in *Mycoplasma,* to 374 of the 470. (Many of the products of the currently unknown genes are likely to be involved in interactions, pathogenic or otherwise, with host organisms.) Many of the known genes can actually be clustered into groups with more narrowly defined missions, but for the sake of this illustration I am grouping them in the more general categories. Assignments of genes to these categories are of necessity incomplete, and in some instances somewhat arbitrary, but I hope they are consistent.

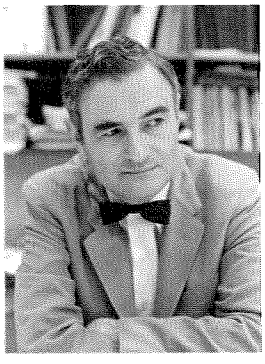
One important additional function that was not considered in 1967 is the provision of means for defense—for survival in a dangerous world, for potential encounters with toxic chemicals, invading viruses, and other cell-eating cells. Both species possess genes for at least rudimentary defense and counter-force measures.

A comparison of the gene distributions for the two microbial species indicates the economies made possible by the parasitic lifestyle of *Mycoplasma genitalium*. Reliant upon host cells for many complex nutrients, this species has a greatly reduced biosynthetic (production) machinery; stealing energy from the host permits much simplification of the energy-producing apparatus; adaptation to life in the relatively controlled environment of a host cell permits great simplification of the means of border control and of the adaptive systems for regulation, control, and repair. Defense agents are trimmed deeply, and even the machinery for information storage and expression is reduced to an essential minimum.

As we unveil the "secrets" of life even in these simple cells and thus define and measure their complexity, this understanding only deepens our marvel at their very existence.

		Haemophilus influenzae	*Mycoplasma genitalium*
1.	cellular envelope, surface structures transport	223 (12.5%)	57 (12.1%)
2.	metabolic reactions	246 (14.1%)	38 (8.1%)
3.	provision of energy	105 (6.0%)	31 (6.6%)
4.	DNA replication, cell division	111 (6.4%)	37 (7.9%)
5.	information transfer	147 (8.4%)	101 (21.5%)
6.	regulation, control, repair	94 (3.6%)	31 (6.6%)
7.	survival, defense	44 (2.5%)	3 (0.64%)

An interesting comparison emerges from the allocation of the genes of two newly sequenced microorganisms into categories devised 30 years ago for defining a living cell. At right, Robert Sinsheimer during his Caltech years.



Robert Sinsheimer was professor of biophysics at Caltech from 1957 and chairman of the Division of Biology from 1968 until leaving to become chancellor of UC Santa Cruz in 1977. He is currently an emeritus professor in the Department of Biological Sciences at UC Santa Barbara.