

Preface

This book introduces the general reader and the specialist to the new order of things in evolution, the origin of life on Earth, and the question of life on Mars and Europa and elsewhere in the universe. Although there are many fields of biology that are essentially descriptive, with the application of information theory, theoretical biology can now take its place with theoretical physics without apology. Thus biology has become a quantitative and computational science as George Gamow (1904–68) suggested. By employing information theory, comparisons between the genetics of organisms can now be made quantitatively with the same accuracy that is typical of astronomy, physics, and chemistry.

Spacecraft send messages to Earth as they pass the outer planets – Mars, Jupiter, Saturn, Uranus, Neptune, and Pluto – in spite of the small amount of energy available. Enormous amounts of data and information flow about on the Internet. Huge sums of money are transferred every day. Errors in these communications cannot be tolerated. Claude Shannon (1916–2001) showed that this is accomplished because communication is *segregated, linear, and digital* so that sufficient redundancy can be introduced in communication codes to overcome errors. Furthermore, he showed that these signals, which contain messages, can be measured in *bits* and *bytes*, terms that are familiar to computer users.

Watson and Crick discovered that there is a genetic message, recorded in the digital sequence of nucleotides in DNA, that controls the formation of protein and of course all biological processes. The message in the genetic information system is *segregated, linear, and digital* and can be measured in *bits* and *bytes*. Computer users will notice the isomorphism between the

program in computer memories and the genetic message recorded in DNA (Chaitin, 1979).

The genetic information system is essentially a digital data recording and processing system. The fundamental axiom in genetics and molecular biology, which justifies the application of Shannon's information and coding theory, is the *sequence hypothesis* and the *digital* rather than the *analog* or *blending* character (Jenkin, 1867) of inheritance as Darwin (1809–82) and his contemporaries believed (Fisher, 1930).

Watson and Crick's solution of the structure of DNA and its application in biology would not have been so important if it had not been for their famously coy remark:

It has not escaped our notice that the specific pairing we have postulated immediately suggests a copying mechanism for the genetic material. (Nature 1953)

A fundamental question in genetics is, how does the cell divide into two cells both containing the same genetics? Here, at one stroke, was the solution nicely framed by reductionism!

I show in this book that only because the genetic message is *segregated*, *linear*, and *digital* can it be transmitted from the origin of life to all present organisms and will be transmitted to all future life. This establishes Darwin's theory of evolution as firmly as any in science. The same genetic code, the same DNA, the same amino acids, and the genetic message unite all organisms, independent of morphology.

The genetic message recorded in the DNA of every living organism is unique to that individual. The relationship and evolution among animals and plants can now be determined by comparing DNA sequences rather than relying on morphology. Genetic information is being applied to genomic medical practice and genetic counseling for the benefit of patients. Sickle-cell anemia is a blood disorder that is an important example of the role of DNA in the placement of amino acids in the sequences of amino acids that form hemoglobin. It is so named because the red blood cells that are normally round are shaped like a sickle. Hemoglobin is composed of four chains of amino acids. Fundamental to this disease, at site 6 in the β chain, glutamic acid is replaced by valine. The identification of this genetic disorder in the structure of hemoglobin with the symptoms of sickle anemia was made by Linus Pauling (1901–94) and is one of his more important discoveries (Pauling, 1949).

DNA now plays a role in forensics identification that is far more important than fingerprints. Forensics has reached new levels of certainty. A number of guilty people have been convicted, and others, falsely accused by conventional methods, have been vindicated.

This is a monograph and not an encyclopedia so I have not considered it necessary to call attention to papers I believe did not make an important contribution or those which are incorrect. I have included in the references only those I felt contributed to the point I was making. Some readers may think that I have neglected an important paper here and there. I acknowledge that this may be the case, but there are times when one must hew to the line and let the chips fall where they may.

This monograph follows my interest in the subject, which was first attracted by the work of Dr. Henry Quastler (1908–63). With his collaboration I organized the *Symposium on Information Theory in Biology* at Gatlinburg, Tennessee, in October 1956. I am indebted to the late Professor Thomas Hughes Jukes (1906–99), whose strong recommendations resulted in my original papers being published. Many of Professor Jukes' important contributions to molecular biology have shaped the ideas presented in this book, particularly those concerning the evolution of the genetic code. I am grateful to Dr. Gregory J. Chaitin, whose original and seminal work in algorithmic information theory is reflected throughout the book. I appreciate the efforts of Dr. David Abel and Mr. John Tomlinson, who read the manuscript and made important corrections and comments. My daughter, Cynthia Ann Yockey, edited this manuscript from proposal to final draft and contributed much to improve the clarity and organization of the material. My editor at Cambridge University Press, Dr. Katrina Halliday, organized the review of the manuscript and arranged for the publication. I appreciate her patience during the writing. Without the contribution of these people I would not have been able to write this book. The reference material is up to date as of February 20, 2004.

Hubert P. Yockey
Bel Air, Maryland, USA