

# An Overview on Synergy between Mathematics and Biology

**Matthew He**

Division of Math, Science, and Technology  
Nova Southeastern University  
Ft. Lauderdale, Florida, USA

E-mail: [hem@nova.edu](mailto:hem@nova.edu)

**Abstract.** Recent progress in the determination of genomic sequences has yielded many millions of gene sequences. But what do these sequences tell us and what are the generalities and rules that are governed by them? It seems that we understand very little about genetic contexts required to “read” them. There is more to life than the genomic blueprint of each organism. Life functions within the natural laws that we know and the ones we do not know. Mathematics can be used to understand life from the molecular to the biosphere level. This paper provides a brief overview of major historical events of molecular biology and genetics, current interface of emerging field of bioinformatics, and future challenges and perspectives between mathematics and biology.

## 1. Introduction

Traditionally the study of biology is from morphology to cytology and then to the atomic and molecular level, from physiology to microscopic regulation, and from phenotype to genotype. The recent development of bioinformatics begins with the research on genes and moves to molecular sequence, then to molecular conformation, from structure to function, from systems biology to network biology and further investigates the interactions and relationships between genes, proteins, and structures. This new reverse paradigm sets the starting point of a biological investigation as theoretical one. It sets a new line of investigation with a unifying principle and extensively uses mathematical tools to quantitatively and analytically clarify the ever-changing phenomena of life.

It's well known there is more to life than the genomic blueprint of each organism [16]. Life functions within the natural laws that we know and the ones that we do not know. Life is founded on mathematical pattern of the physical world. Genetics exploits and organizes these patterns. Mathematical regularities are exploited by the organic world at every level of form, structure, patterns, behavior, interaction, and evolution [2, 16]. Essentially all knowledge is intrinsically unified and lies on a small number of natural laws [9]. Mathematics helps us understand how monomers become polymers necessary for the assembly of cells. Mathematics can be used to understand life from the molecular to the biosphere level including the origin and evolution of organisms [6, 12], the nature of the genomic blueprints and the universal genetic code as well as ecological relationships. Here we briefly walk through the history of interaction between molecular biology and mathematics.



**Mendel's genetic experiments and laws of heredity:** The discovery of genetic inheritance by Gregor Mendel back in 1865 was considered as the start of genetics history [8]. He did experiments on the cross-fertilization of different colors of the same species. The genetic experiments Mendel did with pea plants and took him eight years (1856-1863). During this time, Mendel grew over 10,000 pea plants, keeping track of progeny number and type. He carefully recorded the data and performed mathematical analysis of the data. Mendel illustrated the process that the inheritance of traits could be more easily explained if it was controlled by factors passed down from generation to generation. He concluded that genes come in pairs. Genes are inherited as distinct units, one from each parent. He also recorded the segregation of parental genes and their appearance in the offspring as dominant or recessive traits. He published his results in 1865. He recognized the mathematical patterns of inheritance from one generation to the next. Mendel's Laws of Heredity are usually stated as following:

- **The Law of Segregation:** a gene pair defines each inherited trait. Parental genes are randomly separated to the sex cells so that sex cells contain only one gene of the pair. Offspring therefore inherit one genetic allele from each parent.
- **The Law of Independent Assortment:** Genes for different traits are sorted from one another in such a way that the inheritance of one trait is not dependent on the inheritance of another.
- **The Law of Dominance:** An organism with alternate forms of a gene will express the form that is dominant.

In 1900, Mendel's work was independently rediscovered by DeVries, Correns, and Tschermak. Each of them confirmed Mendel's discoveries. Mendel's own method of research is based on the identification of significant variables, isolating their effects, measuring these meticulously, and eventually subjecting the resulting data to mathematical analysis. Thus his work is directly informed by contemporary theories of mathematics, statistics and physics.

**Origin of species:** Charles Darwin [5] published "On the Origin of Species by Means of Natural Selection" or "The Preservation of Favored Races in the Struggle for Life" in 1895. His key work was that evolution occurs through the selection of inherent and is transmissible rather than acquired characteristics between individual members of a species. Darwin's landmark theory did not specify the means by which characteristics are inherited. The mechanism of heredity had not been determined at that time.

**First genetic map:** In 1910, after the rediscovery of the Mendel's work, Thomas Hunt Morgan [1] did crossing experiments with the fruit fly (*Drosophila Melanogaster*) at the Columbia University. He proved that the genes responsible for the appearance of a specific phenotype were located on chromosomes. He also found that genes on the same chromosome do not always assort independently. Furthermore he suggested that the strength of linkage between genes depended on the distance between them on the chromosome. That is, the nearer two genes lie on a chromosome, the greater the chance of being inherited together. Likewise the farther away they are from each other, the more chance of being separated by the process of crossing-over. The genes are separated when a crossover takes place in the distance in between the two genes during cell division. Morgan's experiments also lead to *Drosophila*'s unusual position as one of the best-studied organisms and most useful tools in genetic research to this day. In 1911, Alfred Sturtevant, then an undergraduate researcher in the laboratory of Thomas Hunt Morgan mapped the locations of the fruit fly genes. This was the first genetic map ever made.

**Transposable genetic elements:** In 1944 Barbara McClintock [13] discovered that genes can move on a chromosome. Genes can jump from one chromosome to another. At that time she studied the inheritance of color and pigment distribution in corn kernels at the Carnegie Institution Department of Genetics in Cold Spring Harbor, New York. At age 81 she was awarded a 1983 Nobel Prize. It is believed that transposons may be linked to some genetic

disorders such as hemophilia, leukemia, and breast cancer. And transposons may have played a crucial role in evolution.

**DNA double helix:** In 1953, James Watson and Francis Crick [17] proposed a double helix model of DNA. DNA is made of three basic components, a sugar, an acid and an organic “base”. The base always was one of the four nucleotides: adenine (A), cytosine (C), guanine (G) or thymine (T). These four different bases are categorized in two groups: purines (adenine and guanine) and pyrimidines (thymine and cytosine). Back in 1950 Erwin Chargaff found that in DNA the amounts of adenine (A) and thymine (T) are about the same, as are the amounts of guanine (G) and cytosine (C). These relationships became known later as “Chargaff’s Rules” and led to much speculation about the three dimensional structure DNA would have. Rosalind Franklin, a British chemist, used the X-ray diffraction technique to capture the first high-quality images of the DNA molecule. Franklin’s colleague Maurice Wilkins showed the pictures to James Watson, an American zoologist, who had been working with Francis Crick, a British biophysicist, on the structure of the DNA molecule. These pictures gave Watson and Crick enough information to propose the double-stranded, helical, complementary, anti-parallel model for DNA in 1953. Crick, Watson, and Wilkins shared the Nobel Prize in medicine and physiology for the discovery that the DNA molecule has a double-helical structure in 1962. Rosalind Franklin, whose images of DNA helped lead to the discovery, died of cancer in 1958 and, under Nobel rules, was not eligible for the prize. In 1957 Francis Crick and George Gamov worked out the “Central Dogma”, explaining how DNA functions to make protein. Their “sequence hypothesis” posited that the DNA sequence specifies the amino acid sequence in a protein. They also suggested that genetic information flows only in one direction, from DNA to messenger RNA to protein, the central concept of the central dogma.

**Genetic code:** The genetic code was finally “cracked” in 1966. Marshall Nirenberg, Heinrich Matthaei and Severo Ochoa [11] demonstrated that a sequence of three nucleotide bases, a codon or triplet, determines each of the 20 amino acids found in nature. This means that there are 64 possible combinations ( $4^3=64$ ) for 20 amino acids. They formed synthetic mRNA by mixing the nucleotides of RNA with a special enzyme called polynucleotide phosphorylase. This resulted in the formation of a single-stranded RNA in this reaction. The question was how this 64 genetic codes could code for the twenty different amino acids, Nirenberg and Matthaei synthesized poly(U) by reacting only uracil nucleotides with the RNA synthesizing enzyme, producing -UUUU-. They mixed this poly(U) with the protein synthesizing machinery of *E. Coli* in vitro and observed the formation of a protein. This protein turned out to be a polypeptide of phenylalanine. They showed a triplet of uracil must code for phenylalanine. Philip Leder and Nirenberg found an even better experimental protocol to solve this fundamental problem. By 1965 the genetic code was almost completely solved. They found that the “extra” codons are merely are redundant: Some amino acids have one or two codons, some have four, and some have six. Three codons (called stop codons) serve as stop signs for RNA synthesizing proteins.

**First recombinant DNA molecules:** In 1972, Paul Berg of Stanford University (USA) [9] created the first recombinant DNA molecules by combining the DNA of two different organisms. He used a restriction enzyme to isolate a gene from a human-cancer-causing monkey virus. Then, he used lipase to join the section of virus DNA with a molecule of DNA from the bacterial virus lambda, creating the first recombinant DNA molecule. He realized the risks of his experiment and temporarily terminated it before the recombinant DNA molecule was added to *E. coli*, where it would have been quickly reproduced. He proposed a one-year moratorium on recombinant DNA studies while safety issues were addressed. Berg later resumed his studies into recombinant DNA techniques, and was awarded the 1980 Nobel Prize in chemistry. His experiments paved the road to the field of genetic engineering and the modern biotechnology industry.

**DNA sequencing and database:** In early 1974, Frederick Sanger [15] from U.K. Medical Research Council was first invented the DNA sequencing techniques. During his experiments

to uncover the amino acids in bovine insulin, he developed the basics of modern sequencing methods. Sanger's approach involved copying DNA strands which would show the location of the nucleotides in the strands. To apply Sanger's approach, scientists had to analyze the composite collections of DNA pieces detected from four test tubes, one for each of the nucleotides found in DNA (Adenosine, Cytosine, Thimidine, Guanine). Then they needed to be arranged in the correct order. This technique is very slow and tedious. It takes many years to sequence only a few million letters in a string of DNA. Almost simultaneously, the American scientists Alan Maxam and Walter Gilbert were creating a different method called the cleavage method. The base for virtually all DNA sequencing was the dideoxy-chain terminating reaction developed by Sanger.

In 1980 Kary Mullis invented polymerase chain reaction (PCR) [3]. The PCR is a method for multiplying DNA sequences in vitro. The purpose of the PCR is to make a huge number of copies of a specific DNA fragment, a gene for instance. The use of thermostable polymerase allows the dissociation of newly formed complimentary DNA and subsequent annealing or hybridization of the primers to the target sequence with minimal loss of enzymatic activity. PCR can be necessary to receive enough starting template for instance sequencing.

Established in 1988 as a national resource for molecular biology information, the National Center for Biotechnology Information (NCBI) carries out diverse responsibilities. NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. NCBI conducts research on fundamental biomedical problems at the molecular level using mathematical and computational methods.

**Human Genome Project:** In 1990, the U.S. Human Genome Project [10] started as a 15-year effort coordinated by the U.S. Department of Energy and the National Institutes of Health. The project originally was planned to last 15 years, but rapid technological advances have accelerated the expected completion date to 2003. Project goals were to:

- identify all the genes in human DNA,
- determine the sequences of the 3 billion chemical base pairs that make up human DNA, store this information in databases,
- improve tools for data analysis,
- transfer related technologies to the private sector, and
- address the ethical, legal, and social issues (ELSI) that may arise from the project.

In 1991, working with Nobel laureate Hamilton Smith, Venter's genomic research project (TIGR) created the 'shotgunning' method. A method at first was controversy among Venter's colleagues who called it crude and inaccurate. However, Venter cross-checked his results by sequencing the genes in both directions, achieving a level of accuracy that greatly impressed his initial sceptical rivals. Within a year, TIGR published the entire genome of *Haemophilus influenzae*, a bacterium with nearly 2 million nucleotides. The draft human genome sequence was published on February 15th 2001, in the journals *Nature* (publically funded Human Genome Project) and *Science* (Craig Venter's firm Celera).

## 2. Living Organisms and Mathematics

Mathematics has a fundamental role in understanding the complexities of living organisms. For example, the genetic code-triplets of three bases in messenger ribonucleic acid (mRNA) that encode for specific amino acids during the translation process have some interesting and mathematical logic in their organization [4]. An examination of this logical organization may allow us to better understand the logical assembly of the genetic code and life.

On the first stage there was investigation on so-called “standard genetic code”. In the last decades some other variants of the genetic code were revealed which are described in a website <http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi> and which differ from the standard genetic code in some correspondences among 64 triplets, 20 amino acids and stop-codons. One noticeable feature of the genetic code is that some amino acids are encoded by several different but related base codons or triplets. There are 64 triplets or codons. In the case of the standard genetic code three triplets (UAA, UAG, and UGA) are nonsense codons - no amino acid corresponds to their code. The remaining 61 codons represent 20 different amino acids. The genetic code is encoded in combinations of the four nucleotides found in DNA and then RNA. There are 16 possible combinations (42) of the four nucleotides of nucleotide pairs. This would not be sufficient to code for 20 amino acids [11]. The solution is mathematically simple. During the self-assembly and evolution of life, a codon or triplet evolved that provides for 64 or 43 possible combinations. This simple code determines all the proteins necessary for life.

The genetic code is also degenerate. For example, up to 6 different codons are available for some amino acid. Another noteworthy aspect of biological messages is that minimal information is necessary to encode the messages [11] and the messages can be encoded and decoded and put to work in amazingly short periods of time. A bacterial *Escherichia coli* cell can grow and divide in 1/2 an hour depending on the growth conditions. Mathematically, it could not be simpler.

Mathematics as applied to understanding biology has countless uses. It is used to elucidate trends, patterns, connections, and relationships in a quantitative manner that can lead to important discoveries in biology. How can math be used to understand living organisms? One way to explore this relationship is to use examples from the bacterial world. The reader is also referred to an excellent text by Stewart [16] that illustrates how math can be used to elucidate a fuller understanding of the natural world. For example, the exponential growth of bacterial cells is essential information that is one of the foundations of microbiology research. Exponential growth over known periods of time is essential in the understanding of bacterial growth in countless areas of research. The ability to use math to describe growth per units of time is an excellent example of the interrelationship between math and the capability to understand this aspect of life. For example, the basic unit of life is the cell, an entity of one. Bacteria also multiply by dividing. Remember that life is composed of matter and matter is composed of atoms and those atoms especially in solids are arranged in an efficient manner into molecules that minimize the energy needed to take on specific configurations. Often these arrangements or configurations are repeating units of monomers that make up polymers. Stewart [16] described it very well in his excellent book when he posed the question. “What could be more mathematical than DNA?” The ability of DNA to replicate itself exactly and at the same time change ever so slightly allows evolutionary changes to occur. The mathematical sequences of 4 different bases (adenine, thymine, guanine and cytosine) in DNA are the blueprint of life. Again, the order of the 4 bases determines the mRNA (messenger ribonucleic acid) sequence and then the protein that is synthesized. DNA in a cell is also capable of precisely replicating itself in a cell. The replicated DNA can then partition into each new cell when one cell divides and becomes 2 cells. The DNA can only replicate with the assistance of enzymes that unwind the DNA and allows the DNA strands to act as templates for the synthesis of the second strand. The ability of a cell to unwind its DNA, replicate or copy new strands and then partition them between 2 new cells has a mathematical basis. The four bases are paired in a specific manner - A (adenine) with T (thymine), C (cytosine) with G (guanine) on the opposite strands along a sugar phosphate backbone. Each strand can contain all 4 bases in any order. However, A must bond with T and C with G on opposite strands. This precise mathematical pairing must be obeyed [4].

Living organisms also have amazing mathematical order and symmetry. The repeating units of fatty acids, glycerol and phosphate that make up a phospholipid membrane bilayer

are one example. An excellent example of mathematical symmetry is the S-layer in many Archaea bacterial (prokaryotes consisting of methanogens, most extreme halophiles and hyperthermophiles and Thermoplasma) cell walls that exhibit a hexagonal configuration. A cell that can assemble the same repeating units countless times is efficient and reduces the numbers of errors incorporated into the assembly. This is exactly the characteristic that is needed for a living cell to grow and divide. Yet, a little bit of change can occur over time.

From a chaotic lifeless environment on the early Earth, life self-assembled with the cell as the basic unit, with mathematically precise order, symmetry and base pairing in DNA as the genetic blueprint and with triplet codons as the genetic code for protein synthesis. Math can be used to understand life from the molecular to the biosphere level. For example, this includes the origin and evolution of organisms, the nature of the genomic blueprints and the universal genetic code as well as ecological relationships. Math helps us look for trends, patterns and relationships that may or may not be obvious to scientists. Math allows us to describe the dimensions of genes, sizes of organelles, cells, organs and whole organisms. Without this knowledge, a paucity of information would still exist on many aspects of life.

### **3. Converting Biological Information to Knowledge**

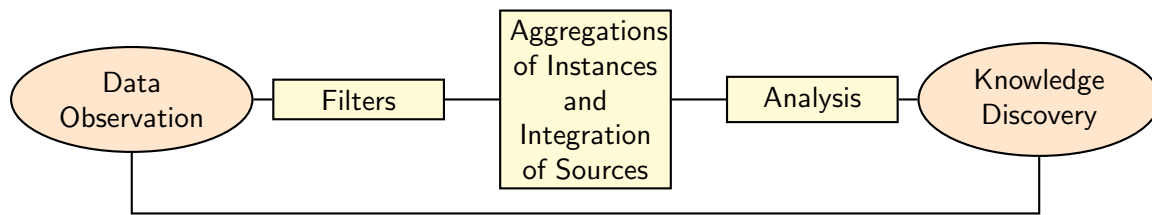
The biological information we gain allows us to learn about ourselves, about our origins, and about our place in the world. We have learned that we are quantitatively strongly related to other primates, mice, zebrafish, fruit flies, roundworms, and even yeast. The findings should induce some modesty, by learning seeing how much we share with all living organisms. The information we are gaining is not just of philosophical interest, but also intended to help humanity to lead healthy lives. Knowledge about primitive organisms provides much information about shared metabolic features, and hints about diseases that affect humans in an economical and ethically acceptable manner.

Knowledge from many scientific disciplines and their subfields has to be integrated to achieve the goals of bioinformatics. It was believed [18] that all knowledge is intrinsically unified, and that behind disciplines, as diverse as physics and biology, anthropology and the arts, lies a small number of natural laws. Applying the knowledge can lead to new scientific methods, to new diagnostics and to new therapeutics.

At the beginning of the “genomic revolution”, a bioinformatics concern was the creation and maintenance of a database to store biological information, such as nucleotide, amino acid, and protein sequences. Development of this type of database involved not only design issues but the development of complex interfaces whereby researchers could both access existing data as well as submit new or revised data. Ultimately, all of this information must be combined to form a comprehensive picture of normal cellular activities. Therefore, the field of bioinformatics has evolved such that the most pressing task now involves the analysis and interpretation of various types of data, including nucleotide, amino acid sequences, protein domains, and protein structures and interactions. Important research branches within bioinformatics include the development and implementation of tools that enable efficient access to, and use and management of, various types of information and new algorithms and statistics with which to assess relationships among members of large data sets, such as methods to locate a gene within a sequence, predict protein structure and/or function, and cluster protein sequences into families of related sequences. The process of converting data to knowledge may be illustrated as in Figure 1.

### **4. Big Pictures: Informatics**

Informatics is the study of the structure, behaviors, and interactions of natural and artificial computational systems. Informatics studies the representation, processing, and communication of information in natural and artificial systems. It has computational, cognitive and social



**Figure 1.** Process of converting data to knowledge.

aspects. The central notion is the transformation of information - whether by computation or communication, whether by organisms or artifacts. Information building blocks may be conceptually as illustrated by Table 1.

**Table 1.** Information Building Blocks (Monomer to Polymer).

Monomer	Polymer
Nucleotides: • Adenine ( A ) • Cytosine ( C ) • Guanine (G) • Thymine/Uracil( T/U)	DNA: ACTGGTAGCCTTAGA... RNA: ACUGGUAGCCUUAGA...
Amino Acids: • Cysteine (Cys) • Alanine (Ala) • Proline (Pro)	Protein: Met-Cys-Gly-Pro-Pro-Arg...
Letters: $A, B, C \dots$	Words: CAT, GO, FRIEND...
Symbols: 0, 1	Binary Code: 1001011100101...
Monomial: $1, x, x^2, \dots$	Polynomial: $P(x), \dots$
Line: $l_1, l_2, l_3, \dots$	Polygons: triangle, rectangle,...

Understanding informational phenomena such as computation, cognition, and communication enables technological advances. In turn, technological progress prompts scientific enquiry. The science of information and the engineering of information systems develop hand-in-hand. Informatics is the emerging discipline that combines the two. In natural and artificial systems, information is carried at many levels, ranging, for example, from biological molecules and electronic devices through nervous systems and computers and on to societies and large-scale distributed systems. It is characteristic that information carried at higher levels is represented by informational processes at lower levels. Each of these levels is the proper object of study for some discipline of science or engineering. Informatics aims to develop and apply firm theoretical and mathematical foundations for the features that are common to all computational systems.

In its attempts to account for phenomena, science progresses by defining, developing, criticizing and refining new concepts. Informatics is developing its own fundamental concepts of communication, knowledge, data, interaction and information, and relating them to such phenomena as computation, thought, and language.

Informatics has many aspects and encompasses a number of existing academic disciplines - artificial intelligence, cognitive science and computer science. Each takes part of informatics as its natural domain: in broad terms, cognitive science concerns the study of natural systems; computer science concerns the analysis of computation, and design of computing systems; artificial intelligence plays a connecting role, designing systems which emulate those found in nature. Informatics also informs and is informed by other disciplines, such as mathematics, electronics, biology, linguistics and psychology. Thus informatics provides a link between disciplines with their own methodologies and perspectives, bringing together a common scientific paradigm, common engineering methods and a pervasive stimulus from technological development and practical application.

Computational systems, whether natural or artificial, are distinguished by their great complexity, as regards both their internal structure and behavior, and their rich interaction with the environment. Informatics seeks to understand and to construct (or reconstruct) such systems, using analytic, experimental and engineering methodologies. The mixture of observation, theory and practice will vary between natural and artificial systems. In natural systems, the object is to understand the structure and behavior of a given computational system. The theoretical concepts underlying natural systems ultimately are built on observation and are themselves used to predict new observations. For artificial systems, the object is to build a system that performs a given informational function. The theoretical concepts underlying artificial systems are intended to secure their correct and efficient design and operation. Computer language systems have been evolving and communicating with biological data as part of computational systems. The computer languages and their interfaces with various data types are illustrated below Table 2.

**Table 2.** Communications between Computer Languages and Data Types and BioModules\*.

Computer Languages	Design Goals
FORTRAN	Numerical analysis
LISP	Symbolic computation
C	System programming
C++	Objects, speed, compatibility with C
Java	Objects, internet
Perl	System administration
Python	General programming

\* BioModules = Bio + Languages

Informatics provides an enormous range of problems and opportunities. One challenge is to determine how far, and in what circumstances, theories of information processing in artificial devices can be applied to natural systems. A second challenge is to determine how far principles derived from natural systems are applicable to the development of new kinds of artificial systems. A third challenge is to explore the many ways in which artificial information systems can help to solve problems facing mankind and help to improve the quality of life for all living things. One can also consider systems of mixed character; a question of longer term interest may be to what extent it is helpful to maintain the distinction between natural and artificial systems.

## 5. Challenges and Perspectives

The interaction between biology and mathematics has been a rich area of research for more than a century. The interface between them presents challenges and opportunities for both



mathematicians and biologists.

Due to the explosion of biological data with the advent of new technologies that can organize the plethora of data, unique opportunities for research and new challenges have surfaced within the last ten to twenty years. For biology, the possibilities range from the level of the cell and molecule to the biosphere. For mathematics, the potential is great in traditional and non-traditional areas such as statistics and differential equations, knot theory, and topology. Stochastic processes and Markov chains in statistics have their origins in biological questions. Modeling the success (survival) over many generations of a family name led to the development of the subject of branching processes. Furthermore biological applications have stimulated the study of ordinary and partial differential equations, especially regarding problems in chaos, fractal geometry, and bifurcation theory.

Further interactions between mathematics and biology have presented new opportunities and challenges. A number of fundamental mathematical and biological issues cut across all of these challenges.

- How do we incorporate variation among individual units in nonlinear systems and biological systems?
- How do we explain the interactions among phenomena that occur on a wide range of scales and molecular levels, of space, time, and organizational complexity?
- What is the relation between pattern and process both in mathematical and biological systems?

It is in the analysis of these issues that mathematics is most essential and holds the greatest potential. These challenges such as aggregation of components to elucidate the behavior of ensembles, integration across scales, and inverse problems are basic to all sciences, in particular to biological sciences, and a variety of techniques exist to deal with them and to begin to solve the biological problems that generate them. However, the uniqueness of biological systems shaped by evolutionary forces will pose new difficulties, mandate new perspectives, and lead to the development of new mathematics. The excitement of this area of science is already evident, and is sure to grow in the years to come [7].

## References

- [1] Allen G 1978. *Thomas Hunt Morgan: the man and his science* (Princeton University Press)
- [2] Anderson P W 1995 Physics: The opening to complexity *Proc Natl Acad Sci USA* **92** 6653-54
- [3] Bartlett J M S and Stirling D 2003 A Short History of the Polymerase Chain Reaction *PCR Protocols* **226** 3-6.
- [4] Cullman G and Labouygues J M 1984 The mathematical logic of life *Proceedings of the 7th International Conference on the Origins of Life*, ed Dose K, Schwartz A W and Thiemann W H P (The Netherlands: D. Reidel Publishing Company)
- [5] Darwin C 1859 *On the Origin of Species by Means of Natural Selection* (London: John Murray)
- [6] Edelstein-Keshet L 2005 *Mathematical Models in Biology, Classics in Applied Mathematics* (SIAM)
- [7] Hastings A and Palmer M A 2003 A bright future for biologists and mathematicians? *Science* **299** 2003-04
- [8] Henig R M 2009 *The Monk in the Garden: The Lost and Found Genius of Gregor Mendel, the Father of Modern Genetics* (New York: Houghton Mifflin)
- [9] Human Genome Project <http://www.genome.gov/25520302>
- [10] Human Genome Project [http://www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)
- [11] Leavitt S 2004 Deciphering the Genetic Code: Marshall Nirenberg *National Institute of Health (NIH)* Retrieved 2009-10-05
- [12] Murray J D 2003 *Mathematical Biology* (Berlin: Springer Verlag)
- [13] Pasachoff N 2006 *Barbara McClintock, Genius of Genetics* (Enslow Publishers, Inc.)
- [14] Prescott L M, Harley J P and Klein D A 1993 *Microbiology* (USA: Wm. C. Brown Publishers)
- [15] Sanger F, Donelson J E, Coulson A R, Kössel H and Fischer D 1974 Determination of a nucleotide sequence in bacteriophage  $\phi$ 1 DNA by primed synthesis with DNA polymerase *J Mol Biol.* **90** (2) 315-33
- [16] Stewart I 1998 *Life's Other Secret* (USA: John Wiley and Sons, Inc.)

- [17] Watson J D and Crick F H C 1953 A structure for deoxyribose nucleic acid *Nature* **171** 737-738
- [18] Wilson E 1998 *Consilience: The Unity of Knowledge* (Random House)