Fundamentals of DNA Computation in the Domain of Cryptosystems

Adithya B. (f51844e1-5690-4bbc-afd8-22e81de10298, Puducherry Technological University, India
Santhi G. (a711be60-0dcf-43de-a46b-b0de7a16eb7e, Puducherry Technological University, India

ABSTRACT

In the modern world of e-business and e-commerce, security of the confidentiality, integrity, and availability (CIA triad) of stored information along with transmitted data is essential. In cryptography, DNA cryptography is a newly emerging discipline arising from the work of DNA computing. The idea of DNA cryptography is based on DNA molecules capable of storing, processing, and transmitting information. The biological history of DNA cryptography and the theory of DNA computation are briefly described in this article. In comparison to traditional cryptography and quantum cryptography areas, DNA cryptography progresses via security and its applications. The primary objective of DNA cryptography research is to study the DNA molecule's features and reactions, create appropriate hypotheses, discover potential development pathways, find simple DNA cryptography implementation methods, and lay the groundwork for future growth.

KEYWORDS

Cryptography, DNA, DNA Computing, DNA Cryptography, Quantum Cryptography

INTRODUCTION

The immense parallelism, the tremendous power consumption and the exquisite density of data revealed in DNA molecules are being investigated for computing, information storage, and cryptography. In these studies, new computers, information storage and cryptography can be identified, leading to new developments in information technology. On this sort of foundation, DNA cryptography builds on the work of DNA computing (also called organic or molecular computing). With the advent of computer technology in the 20th century, modern cryptography made great strides and is still used widely. DNA cryptography is built on that foundation with the work of DNA computing. Conventional cryptography made great strides with the advent of computer technology in the 20th century and is still widely used. Since Adleman first suggested DNA computing in 1994, DNA cryptography has received much attention and has become cryptography's frontier. Under the same protection of intent information, DNA cryptography, conventional cryptography, and quantum cryptography are constructed differently. All three forms of cryptography may be the key areas of potential cryptography. In this article, the researcher explores and examines the biological history, scientific success, and prospects of DNA cryptography to highlight potential studies.

DOI: 10.4018/IJISSC.322395

*Corresponding Author

BIOLOGICAL ANTECEDENTS

There are two biological antecedents:

1. Deoxyribonucleic Acid (DNA) preliminaries
2. DNA Technologies as follows.

Deoxyribonucleic Acid

DNA is an acronym for Deoxyribonucleic Acid, which in all lifestyles is a plasmatic insect. DNA is a type of biological macromolecule made up of nucleotides. Adenine (A) and Thymine (T) or Cytosine (C) and Guanine (G), which are four nucleotide organisms, each have just one base. This is how single-stranded DNA is built: the 5’ termination is termed 5’, and the 3’ termination is termed 3’. In nature, DNA is usually found as double-stranded molecules. Two complementary DNA strands are kept together by hydrogen bonds formed between the complementary bases of the A and T (or C and G) strands, as shown in Figure 1 (Adithya & Santhi, 2019). Watson and Crick discovered the double helix structure. Watson-Crick complementation (Watson et al., 1987) is the name given to the complementary structure. The researcher’s work, which reduced biology to chemistry and marked the beginning of biology in the second part of the century, is one of the greatest scientific triumphs of the 20th century (Seeman, 2004). As mentioned in Table 1, few scientific biological concepts should work with DNA computing and cryptography.

Figure 1.
Structure of DNA

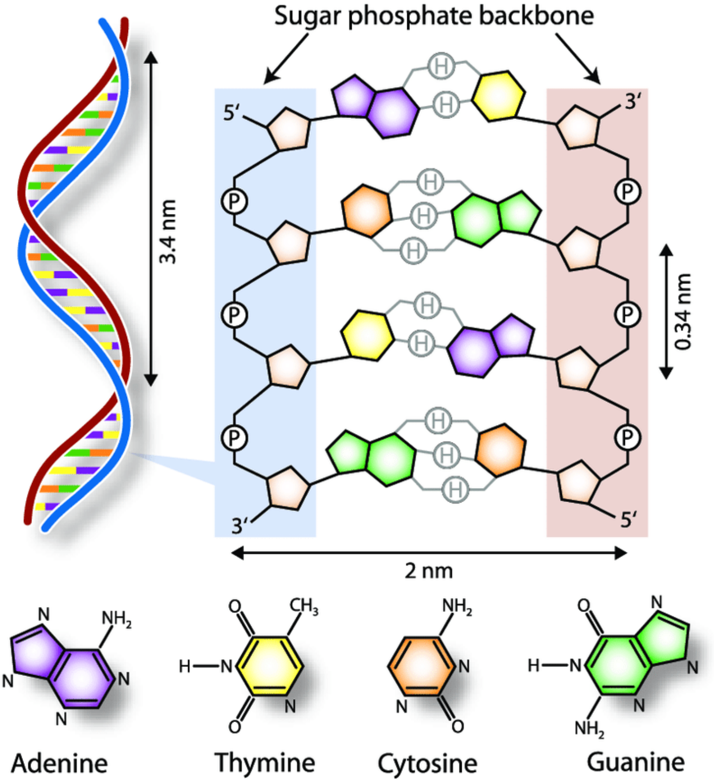


Table 1.
Biological concepts of DNA computing

Basics of DNA	Descriptions
Codon	Codon is a triplet-form sequence of three nucleotide bases.
Gene	Gene is the functioning DNA subunits. Specific set of instructions to be carried by every gene to code a certain protein or functions. It includes both coding DNA series (exons) and non-coding DNA series (introns).
Chromosome	Chromosome is a broad structured DNA structure coiled around proteins, and includes genes, regulatory elements, and other sequences of nucleotides. The chromosomes can consist of long gene strings.
Genome	Genome is an organism's unique sequence that includes a cell's DNA content, nucleotides, proteins, and chromosomes.

DNA Technology

Over the decades, there have been significant advances in genetic engineering. Because those technologies are advanced and matured, consisting of DNA synthesis, DNA sequencing hybridization, electrophoresis, Polymerase Chain Reaction (PCR), and other strategies, scientists can effortlessly expand DNA sequencing parts by using isolated, digestible, and amplified sequential strings. Inside the improvement of existence, this opens the door to new applications of diverse types of DNA. In all current DNA cryptography and DNA computing approaches, PCR technology is used. In contrast, DNA hybridization performs much better than DNA chip technology. Gel electrophoresis technology for separating or extracting DNA fragments in a gel. By cutting long DNA at random places on the chain, many short DNA fragments are produced. The chaotic theory is used primarily for the encryption of images using DNA. This paper briefly introduces these critical technologies.

Primitive Operations on DNA Strands

Primitive operations on DNA strands are:

1. DNA Synthesis
2. DNA Hybridization and Denaturation
3. DNA Cutting and Ligation
4. DNA Separation and Extraction.

DNA Synthesis

Whilst a cell divides, DNA biosynthesis occurs in a method known as replication. It needs to separate the DNA double helix and then use the parent DNA root or strand as a template to synthesize a complementary DNA strand (see Figure 2). Textual content written in any language is encoded in 4 alphabets A, C, G, and T.

DNA Hybridization and Denaturation

Watson-Crick based on complementary theory, single DNA strands of opposite dendrites combine to form a dual helical structure is shown in Figure 3. This method is also referred to as annealing. The opposite technique of annealing is called denaturing.

DNA Cutting and Ligation

It selects a unique quick-length DNA sequence known as the restrict enzyme. The restriction enzyme is located in double-stranded DNA. The site where this restriction endonuclease is produced is called

Figure 2.
DNA synthesis

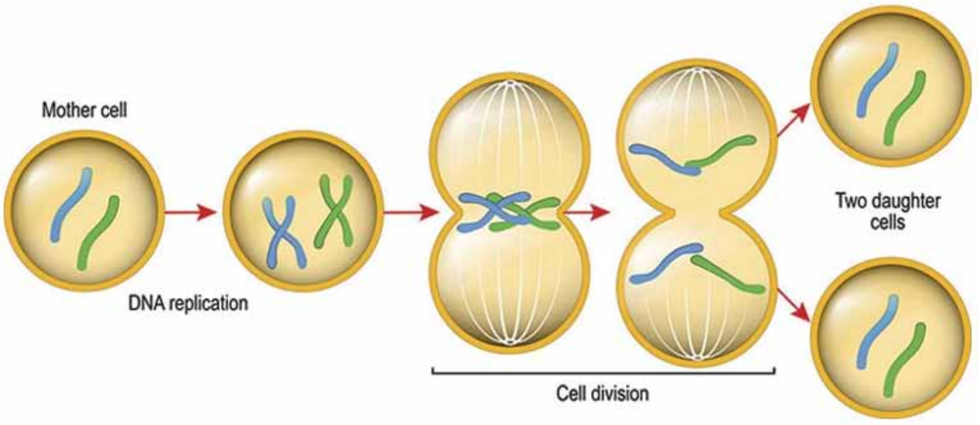
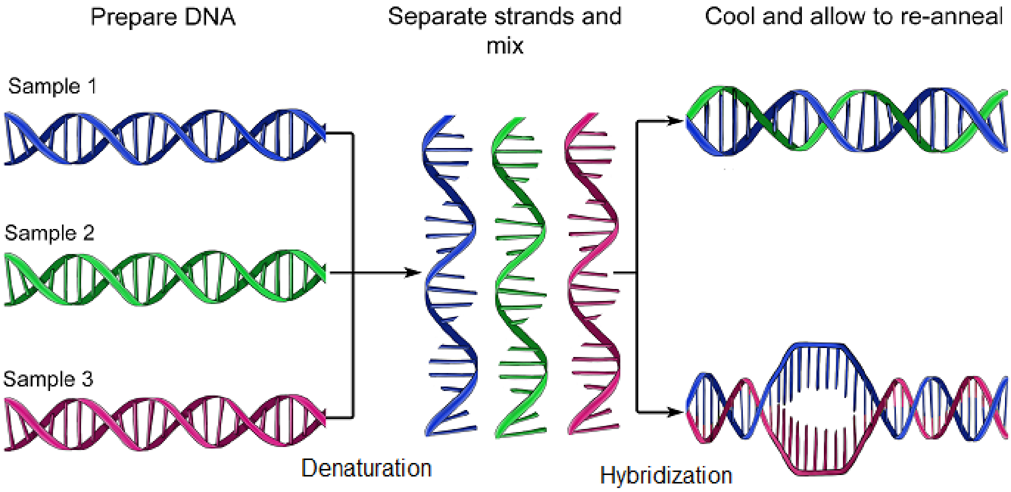


Figure 3.
DNA hybridization and denaturation

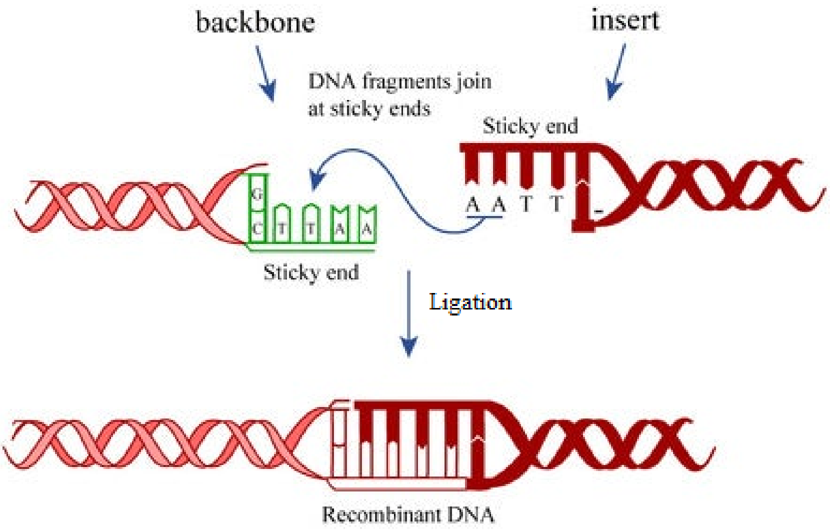


the restriction site and is contained in the double-stranded DNA sequence (Adithya & Santhi, 2021). The enzyme breaks the DNA sequence from a location that is the same as the enzyme in a certain series. As a result shown in Figure 4, two double-stranded DNA with "blunt ends" or two double-stranded DNA with single-stranded cantilevers are formed, referred to as "sticky ends." The reverse cutting process is ligation. In this, the resulting double-stranded DNA cutting activity sequences can be fixed and rejoined by an enzyme known as DNA ligase.

DNA Separation and Extraction

This procedure is performed using the technique of gel electrophoresis. In this, electrophoresis filters out DNA molecules according to their size (small or large). The affinity purification process removes a single-stranded DNA molecule containing the intended base sequence is shown in Figure 5.

Figure 4.
DNA cutting and ligation



Development and Maturation Technologies in DNA Strands

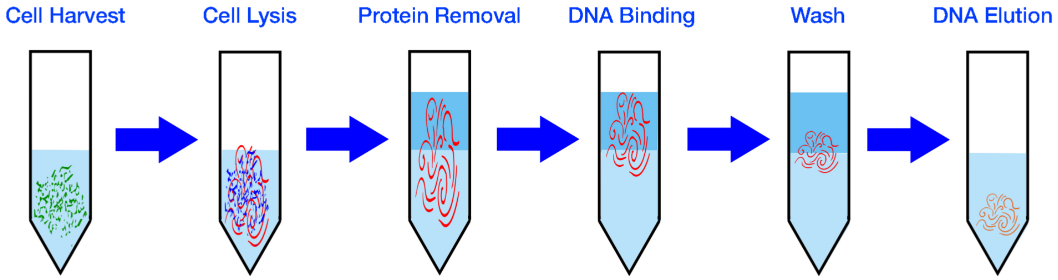
Development and maturation technologies in DNA strands are:

1. Gel Electrophoresis
2. DNA Chip and Polymerase Chain Reaction
3. Fragment Assembly
4. Chaotic Theory as follows.

Gel Electrophoresis

It is a method that is used to distinguish the DNA fragments through their weight. A polyacrylamide gel or agarose gel prepares it. On one side of the gel, molecules of negatively charged DNA are placed into wells. The negatively charged DNA molecules begin to move towards the positive pole when an electrical current is applied to the gel is shown in Figure 6. The lower molecules move faster than the larger ones.

Figure 5.
DNA separation and extraction



Hence, it is easy to detect a distinction between them (Kari, 1997; Lodish et al., 2003; Paun et al., 1998). Strands of the same length travel at the same pace and end up clustered together, thereby sorting themselves the DNA strands in the sample. Finally, it acts as a filter between the longer and shorter bands.

DNA Chip and Polymerase Chain Reaction

DNA chips are also identified as biologic chips made from DNA microarrays or gene chips or oligonucleotide chips or as indicated oligo nucleic acid or cDNA samples based on reported Fodor and Brown methods (Fodor et al., 1991; Pease et al., 1994; Schena et al., 1995; Shalon et al., 1996). Hundreds of millions of DNA probes are stacked on a glass or silicon substrate in the space of less than 1 square inch. And their counterparts use multiple-label samples to reveal genetic information with samples on a chip to obtain various hybrid spectra, as shown in Figure 7. Microarray steps are illustrated in Figure 7(a); in this, gene expression patterns between healthy cells and cancerous cells are compared. Information regarding microarrays can be represented as a heat map. Figure 7(b) shows that the gene is displayed, and researchers can see various specimens below. Genes that are expressed only in cancer cells are offered in a different color, red. Genes that are only expressed in normal cells are shown in a distinct green color. Genes expressed in both cancer cells and normal cells are shown in yellow. Therefore, researchers can increase hybridization efficiency thousands of times.

Researchers discovered the Chain Reaction to Polymerase in 1983. It is one of the most critical innovations in the biology of modern (Debao & Ping, 1994). Since the DNA molecule is tiny in volume, a small amount of DNA provided is difficult to manipulate directly, whereas working a large amount of DNA after amplification would be very easy. PCR is a Watson-Crick-based fast and complementary DNA amplification technology. After a sequence of 5’ to 3’ polymerase reactions, pair two oligonucleotide primers corresponding to the double-stranded target DNA strand and amplify the desired target DNA, as illustrated in Figure 8. PCR is a susceptible technique: after 20 cycles, researchers can theoretically amplify 10^6 single target DNA molecules. Therefore, in a short time, one can potentially strengthen many strands of DNA (Debao & Ping, 1994).

Fragment Assembly

In this approach, the original lengthy DNA strand is recreated from a large variety of DNA fragments. The amplification procedure is then used to increase the DNA. A large variety of short DNA fragments is then acquired by applying the spontaneous dissection of a long DNA fragment, as shown in Figure 9. Then, by precisely combining such DNA fragments, the unique long DNA sequence is obtained so

Figure 6.
Gel electrophoresis

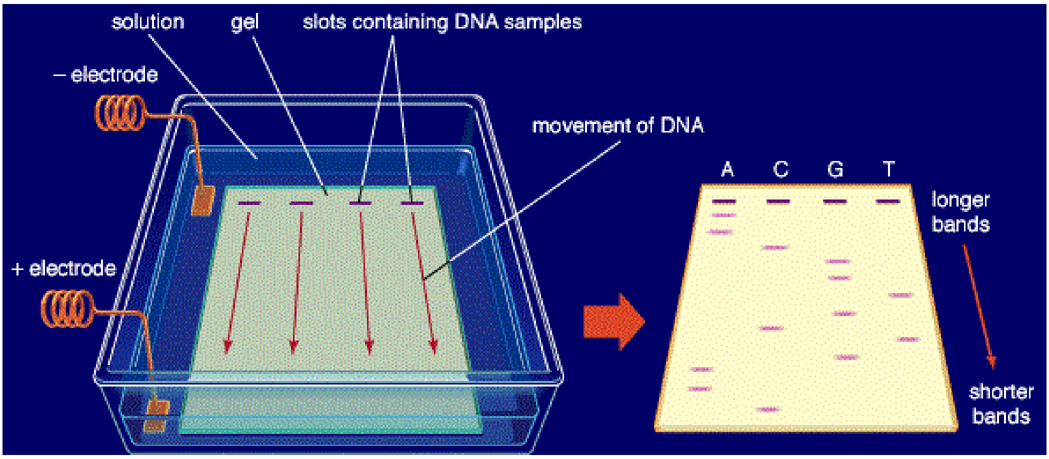
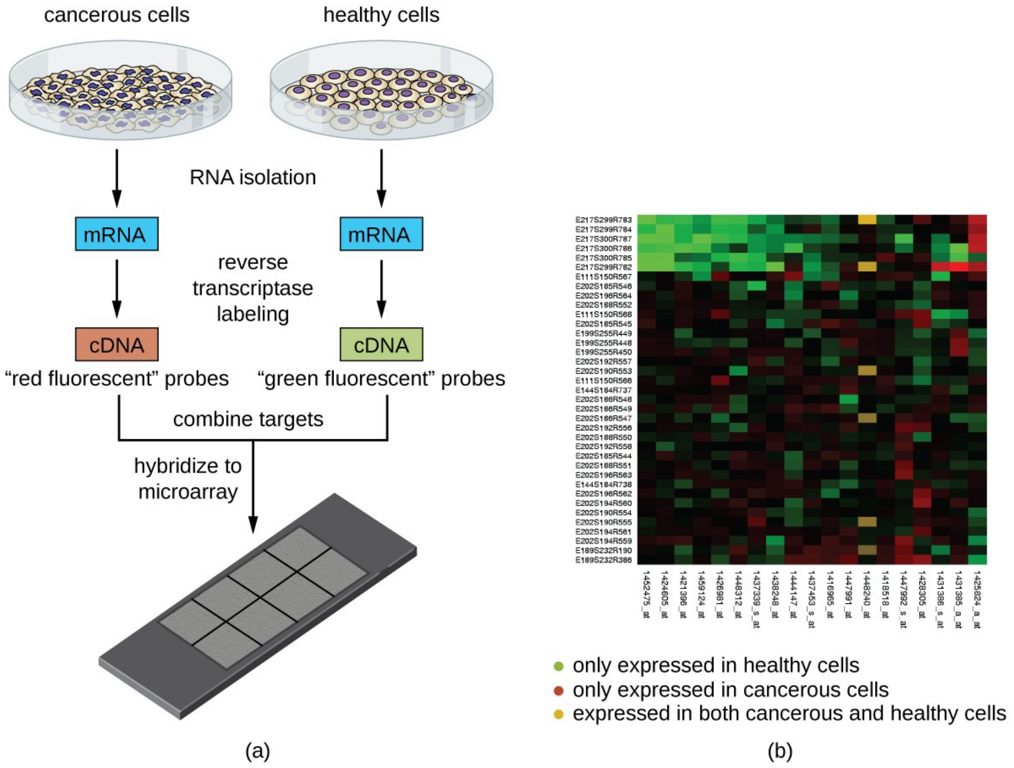


Figure 7.
DNA mircoarray



that the final part of a DNA fragment overlaps with the initial part of the second order. This technique is likewise known as “shotgun sequencing” (Gibson, 2011).

Chaotic Theory

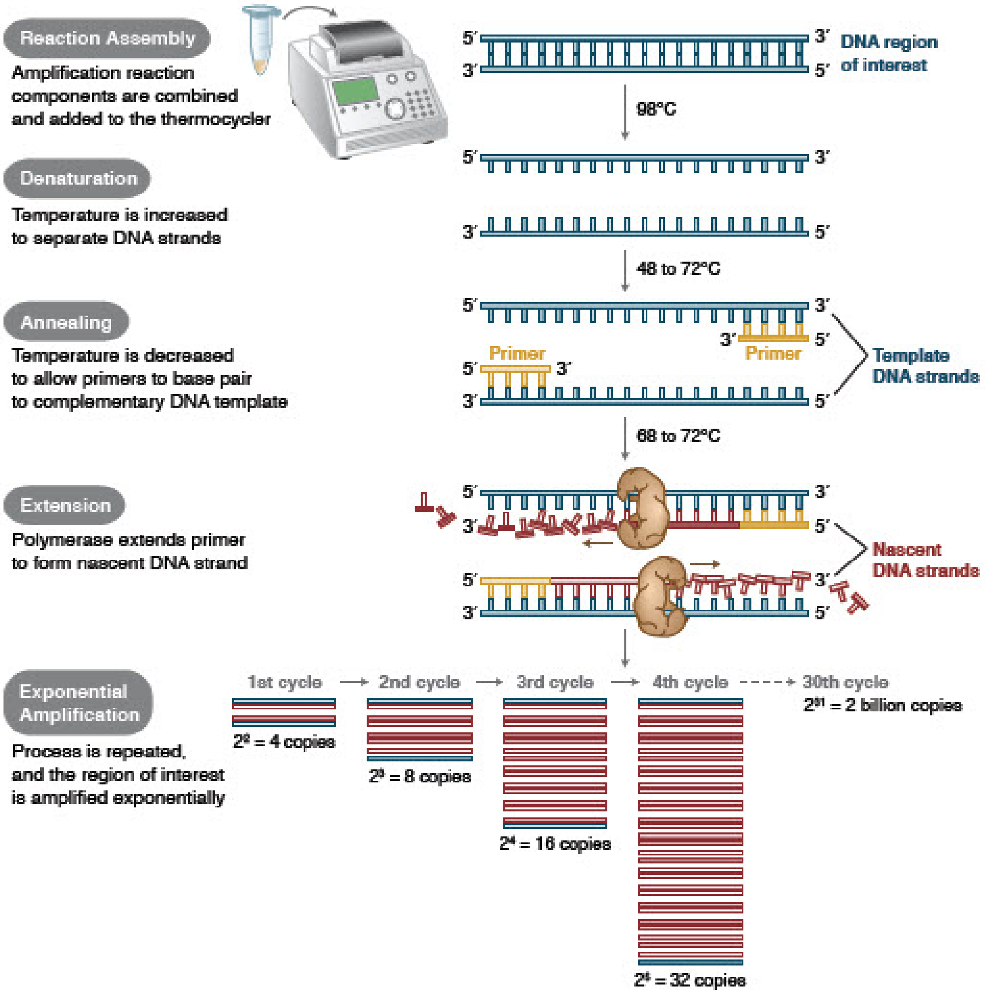
Dynamic processes are susceptible to the original conditions according to chaos theory. Even if the system is suppressive, slight differences in initial conditions can cause widespread confusion in the results. The behavior of such machines is called chaos. This principle is primarily used for image encryption using DNA.

DNA COMPUTATIONS

Advances in DNA computing (also called molecular or organic computing) contribute to the improvement of DNA cryptography. On the one hand, cryptography regularly has extra or less something to do with the corresponding computational version. Then again, DNA cryptography also includes a few organic strategies used inside the computation of DNA. For those purposes, DNA computations are in short applied here. Adleman demonstrated the primary computation of DNA in 1994, which marked the start of a brand new level in statistics technology (Adleman, 1994). Scientists consider the massive parallelism, exceptional energetic capacity, and remarkable density of information found in DNA molecules in the following research (Adleman, 1994; Guarnieri & Fliss, 1996; Ouyang et al., 1997; Sakamoto et al., 2000).

A rudimentary DNA machine led by Adleman solved three SAT queries with more than 1 million results after an intensive search in 2002 (Ravenderjit et al., 2002). In 2005, a crew led by

Figure 8.
Polymerase chain reaction

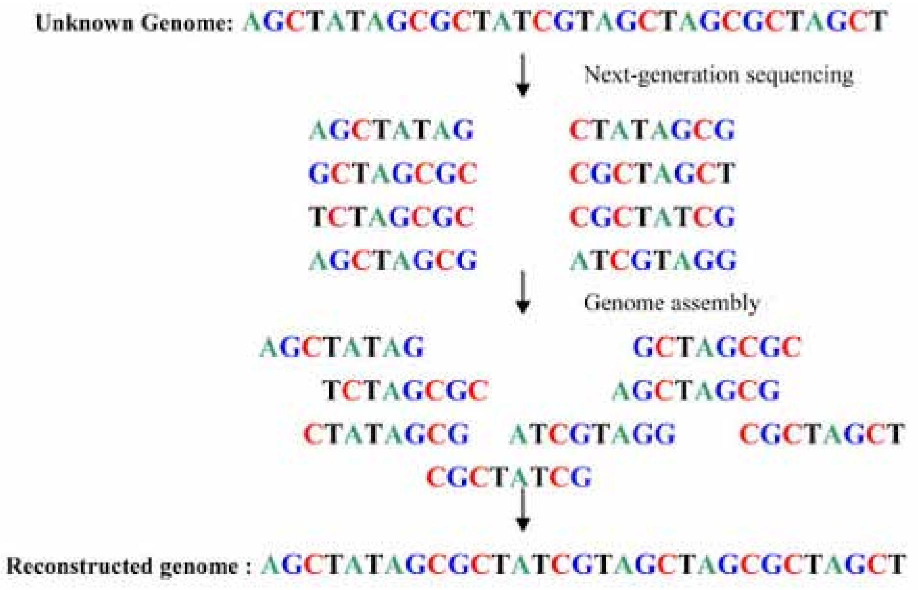

Reaction Assembly
Amplification reaction components are combined and added to the thermocycler
Denaturation
Temperature is increased to separate DNA strands
Annealing
Temperature is decreased to allow primers to base pair to complementary DNA template
Extension
Polymerase extends primer to form nascent DNA strand
Exponential Amplification
Process is repeated, and the region of interest is amplified exponentially
5′
3′
DNA region of interest
98°C
48 to 72°C
Primer
Primer
Template DNA strands
68 to 72°C
Nascent DNA strands
1st cycle
2nd cycle
3rd cycle
4th cycle
30th cycle
2^{31} = 2 billion copies
2^2 = 4 copies
2^3 = 8 copies
2^4 = 16 copies
2^5 = 32 copies

Ehud Keinan carried out a DNA test, and researchers found the biomolecular gadgets to use a couple of that could perform up to a billion operations concurrently. Adleman analyzed DNA computing as follows: “Humans have tried for thousands of years to use engineered tools to improve their innate computational capabilities.” Significant trends were mechanical gear along with the abacus, the including gadget, and the tabulating machine. Scientists have made further advances in the concept of DNA computation and have tested many possible computational models using functions, such as the Adleman model (1994). This case is called the Hamilton path model, DNA chip version (Liu et al., 2000), and Adelman sticker version (Roweis at al., 1998). Below is a brief overview of commonly used Hamilton path patterns and sticker patterns.

Hamiltonian Path Design

Adleman’s (1994) DNA computation is to solve Hamiltonian guideline problem. Lipton (1995) also uses the computing model that Adleman has proposed to solve SAT problems. The challenge of the Hamiltonian method is to find a path that starts with X_{in} and ends with X_{out} and connects all other

Figure 9.
DNA fragment assembly



vertices once on the directed graph. A 20-mer (short DNA strand) random oligonucleotide of ON_i is created for each graph's vertex. Here, the long oligonucleotide measure is mer. The corresponding $ON_2, ON_3,\ and\ ON_4$ represents vertices 2, 3, and 4, respectively. Each oligonucleotide below is written between 5'and 3':

$$ON_2 = ATCGCTAGGTACCGTATTCA$$

$$ON_3 = AGCCTATATGGCTAGGCTAT$$

$$ON_4 = TTCGTACGACCATGGATCGG$$

For each edge $i \rightarrow j$ in the graph, an oligonucleotide strand extracted from ON_i is 3' 10-mer and ON_j is 5' 10-mer. $\overline{ON_i}$ is Watson-Crick complement of each vertex of i in the graph ON_i . Except for $\overline{ON_3}$, all oligonucleotide at the end of Figure 10 are written as 5' to 3':

$$ON_2 \rightarrow_3 = ACCGTATTCAAGCCTATATG$$

$$ON_{34} = GCTAGGCTATTTCGTACGAC$$

$$\overline{ON_3} = TCGGATATACCGATCCGATA$$

In the experiment, ample $\overline{ON_i}$ and $ON_i \rightarrow_j$ was combined in a single ligation reaction for each vertex of *i* (except X_{in} and X_{out}) and for each edge of $i \rightarrow j$ in the graph. A lot of oligonucleotides were created after a "graduated PCR" reaction which denotes different paths. The $\overline{ON_i}$ oligonucleotides were used as splints for ligation combining oligonucleotides associated with compatible edges, as shown in Figure 5. After completing all the reaction phases, to amplify paths that began from X_{in} and ended at X_{out}, a PCR was performed. Then, gel electrophoresis separated oligonucleotides with the right weight. A network of biotin-avidin magnetic beads isolated the paths, which did not reach any vertex. The Hamiltonian route will be presented if there are any leftover DNA strands.

Sticker Design

The Hamiltonian route and the sticker layout are similar to the Watson-Crick complement. The difference is that there are small chains and longer chains in the Hamilton path model that affect the results throughout the annealing process, whereas the sticky tag mode carries long single-stranded DNA first, and the quick tag sticks to a chain. The sample of the stickers is utilized in Adleman et al. (1999), and Ravenderjit et al. (2002). The sticker model contains the memory structure primarily with four operations: isolation, combination, clearing specific bits, and setting specific bits. The sticker model memory structure is shown in Figure 11 (Guozhen et al., 2006). More detailed information about the model sticker is shown in Roweis et al. (1998). Despite advances in DNA computing in theory and practice, there is still a long way to go before all the secrets of the cell can be unraveled. Therefore, in future research, new computational models are likely to be found. As per Giffford (1994), "Transcriptional control and other methods of gene regulation certainly play an important role in the programming of cellular behavior, but there may be other computational methods that involve mostly basic biological processes".

BASIC DNA CRYPTOGRAPHY DEVELOPMENT ISSUES

DNA cryptography is nascent cryptography wherein DNA is used as a records service, and cutting-edge biotechnology is used as an implementation technique. A broad parallel to the records worried in the DNA molecule with excellent power performance and outstanding density in encryption, authentication, signature, and many others. Adelman's proposed DNA calculation can't be used as direct DNA cryptography. Even in studying various complex biological problems in DNA cryptography, DNA generation clarifies complex calculation problems in DNA computing. It is used as a stable basis for the DNA cryptography system. The form of encryption and the decryption process can be seen as a kind of calculation. In cryptography, however, all DNA computations are superfluous. Keep in mind that DNA cryptography is not the same as genetic coding. The genetic code is related to the origin of life, and genes are part of biotechnology. Currently, research work is in the early stage of DNA cryptography, and (Gehani et al., 2000; Celland et al., 1999; Leier et al., 2000) reference is included in some of the most successful schemes of DNA cryptography.

Figure 10.
Directed graph

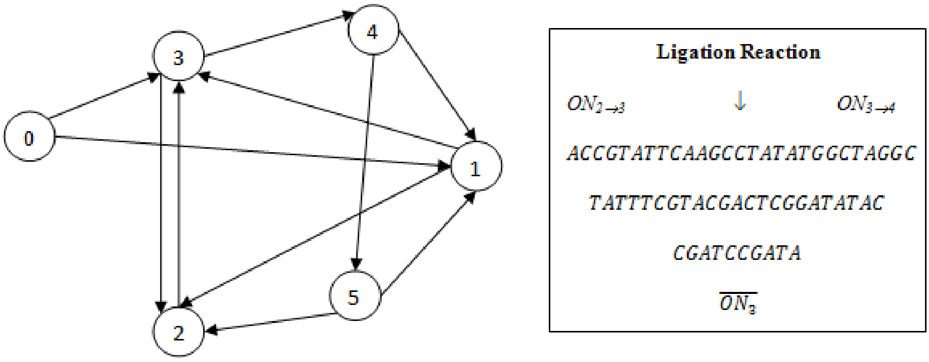
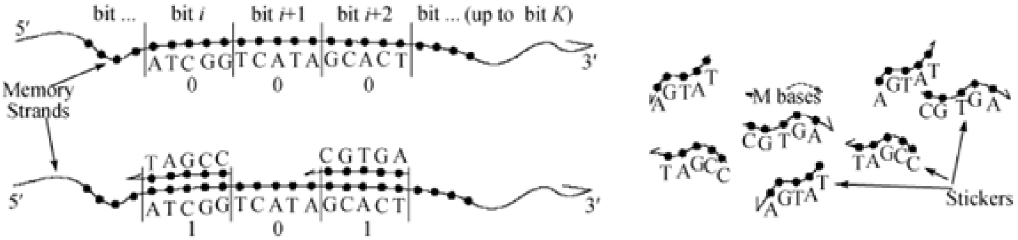


Figure 11.
Sticker design



This article provides the first representative DNA encryption scheme. Initial DNA indicates that the density of facts is high, but this approach can be challenging to implement. When using PCR for decryption, the arrival process is much simpler. The PCR log analysis technique is not only related to the calculation of DNA but has also been widely used in the following investigations. Because of its dependability, the second approach is often known as DNA steganography. It is a good idea to think of it as a DNA cryptography technology.

Gehani et al. (2000) completed a DNA-primarily based one-time pad approach. The current implementation of smart cryptographic devices based on one-time pads, according to researchers, is restricted to the scope of traditional electronic media. However, DNA offers a high recording density. 10^{21} DNA bases or 10^{8} terabytes are contained in one gram of DNA. Therefore, few grams of DNA can save all of the records accumulated around the arena. Thus, DNA is very suitable for storing giant one-time pads. Gehani et al. (2000) proposed scheme is wholly based on excellent DNA data density and has a remarkable efficiency capacity value. This solution is capable of resolving the issue of one-time pad storing. The downside is that preparing a large DNA one-time pad for rapid separation and data reading is difficult. Complex biological tests for the transmitter and receiver are required, which can only be carried out in a well-equipped laboratory and is prohibitively expensive. For the above reasons, the scheme is not feasible for several years.

Clelland et al. (1999) the "June 6 invasion: Normandy" was successfully masked in DNA microbes for DNA steganography. The process will be as follows.

Encodification Statute

Instead of conventional binary encoding, a new encoding approach is proposed. Each letter is represented by three nucleotides, which are employed as a quaternary code. For example, CGA denotes the letter A, CCA distinguishes the letter B, etc.

DNA Secret-Message Synthesizer

The secret message is encoded into the DNA chain according to the code mentioned earlier. The letter AB, for example, is encoded as CCGCCA. Following encoding, a secret DNA oligodeoxynucleotide message is created using forward and reverse PCR primers. Every 20 nucleotides carry 69 nucleotides long-code information. Therefore, the secret DNA information is ready.

Data Hiding

Researchers intend to sonicate human DNA with 50 to 150 nucleotide pairs to hide and denature DNA that is physically comparable to secret information DNA (universal size). The concealed DNA and the sensitive information DNA are combined and then transferred to a paper with a conventional adhesive to form colorless micro-dots. General mail service will then post the article containing microdots.

Decodification Statute

Encoding law and primers are the communication codes for the sender and the receiver. Once you have received the recipient's paper, you can effortlessly search for microdots since the desired recipient had obtained the primer and coding law securely. Researchers could perform the DNA microdot PCR following the coding law, sort and retrieve the secret message (plaintext) to the DNA, could promote amplification.

Global Developmental Issues

Global progress shows that today's DNA cryptography is facing the following problems:

Lack of a Theoretical Framework for DNA Cryptography

In 1949 Shannon proposed in his famous paper, "Communication theory of secrecy systems," the basic model for modern privacy communications and development direction (Shannon, 1949). In the 1970s, complexity theory was proposed to be a powerful method for developing encryption algorithms, which helped create public key cryptosystems (Diffie & Hellman, 1976). New cryptosystems, which include DES, RSA, AES, and EIGamal, had been discovered inside the following a long time (Daemen & Rijmen, 2002; EIGamal, 1985; NIST, 1993; Rivest et al., 1978). Consequently, conventional cryptography is more and more being perfected. In evaluation, no corresponding mature principle exists for DNA cryptography. It is miles nevertheless an open question about the version and safety base of DNA cryptography, nothing to mention about implementation. Because of the unavailability of a relevant idea, it's far tough to make exact mystery DNA schemes.

Hard to Know and Expensive to Implement

In the encryption and decryption phases of the present DNA encryption method, researchers must execute several biological experiments, such as producing the message DNA strand and performing PCR amplification and sequencing. Only a well-equipped laboratory with modern equipment can do such trials. DNA cryptosystems are not feasible for such reasons and cannot compete with conventional cryptosystems. Fortunately, in the last twenty-five years, modern biology has come a long way. Many once-expensive experiments have since become commonplace. The "complex and expensive" challenge has been overcome by advances in biology and contemporary DNA code design.

COMPARISON OF DNA CRYPTOGRAPHY, QUANTUM CRYPTOGRAPHY, AND CONVENTIONAL CRYPTOGRAPHY

Conventional, DNA, and Quantum cryptography is compared with 1) evolution, 2) reliability, and 3) related applications.

Evolution

Conventional cryptography dates again to Caesar encryption 2000 years ago or maybe in advance. The associated principle is almost correct. Traditional cryptography includes all workable ciphers. Quantum cryptography became hooked up inside the 1970s, and a theoretical foundation was organized, even though hard to put into effect. In well-known, researchers did not plunge their findings into practice. DNA cryptography has most effective nearly fifteen years of history, the theoretical foundations are being studied, and its software costs a lot of money.

Reliability

Computational security can simplest be carried out for classic cryptographic projects, except for the one-time pad. i.e., with infinite computing power, the opponent can theoretically break them. Quantum computers have been shown to have enormous and astonishing computational capabilities (Grovel, 1996; Shor, 1994; Simon, 1994). Despite the ambiguity regarding the computational power of quantum computers, the use of future quantum computers is likely to disrupt all conventional circuits except the one-time pad. According to modern theories, quantum cryptographic projects are indestructible.

Nevertheless, the researcher's defense is based on Heinsberg's uncertainty theory. However, suppose the snooper is given the freedom to do whatever he wants and has unlimited computing resources equivalent to P = NP to execute. In that case, such a scheme is still difficult to decipher. Some eavesdropping activity will adjust the encryption so it can be detectable. An adversary cannot obtain a wholly identical quantum with the intercepted one, therefore an effort to exploit but not detected in vain (Wiesner, 1983; Bennett & Brassad, 1984; Bennett, 1992; Ekert, 1991; Bennett et al., 1992). Hence, the schemes for quantum key agreements have absolute security. The necessary protection of DNA cryptography is to comply with organic technology that has nothing to do with computing power and the security of the cryptographic DNA scheme to prevent quantitative code attacks. But, the degree of this shape of protection and the way lengthy could be maintained are still being investigated.

Applications

The traditional cryptosystem is the most convenient. Electronic, quantum and DNA computers can calculate the information. Data can be sent by wires, optical fibres, wireless networks, or even Messenger. It can also save it on CDs, magnetic media, DNA, and other sorts of storage. Researchers may also employ traditional encryption approaches to accomplish goals like public and private key encryption, identity authentication, and digital signatures. The primary benefit of quantum cryptography is real-time communication, which is based on quantum networks. The drawback is the safe data storage, which makes typical public-key encryption and digital signatures hard to enforce. The ciphertext in DNA cryptography can only be conveyed by physical techniques, according to current technological requirements. Due to the considerable parallelism, excellent energy efficiency, researchers could inherent high recording density of DNA molecules.

DNA encryption can have unique benefits in some encryption software packages, including secure data storage, identity verification, virtual signatures, steganography, and many other advantages. In addition, DNA can be used to create hard-coded contracts, cash coupons, and ID cards. Research on all three types of cryptocurrencies is ongoing, and there are many areas to be resolved. Especially for DNA and quantum cryptography, the future is hard to look forward to. But from the previous discussion, researchers have concluded that researchers are likely to exist and evolve with all the differences and complement each other, rather than one of them being deeply abandoned.

DIRECTION OF ESTABLISHING DNA CRYPTOGRAPHY

Given that DNA cryptography is still in its infancy, it is still too early to make accurate predictions about its future development. However, due to the advancement of biological technology and the demands of cryptography, we reserve the following positions.

Security Basis and Resources of DNA Cryptography

Current biological tactics should be used as assets, and complicated organic difficulties should be used as the vital protection basis for DNA cryptography. Encryption and decryption are data transformation approaches described by mathematical techniques that can be less complicated to enforce in the present generation of digital computers and the net rather than physical and chemical ones. Researchers will have more excellent security ratings and storage capacity if specific cryptographic systems need to be explored and created where electronic computers cannot identify the use of mathematical procedures. As a result, if the development of DNA cryptography is appropriate, the benefits inherent in DNA need to be very well explored.

The story of nano-level storage is mainly based on the small area of DNA and entirely relies on significant parallelism for fast encryption and decryption. Using complicated biological problems, it is still far from being completely understood as the stable DNA cryptography foundation to realize a new cryptosystem that can withstand the attack from quantum computers. Because it has not been clearly stated that quantum computers are threatening the hardness of different mathematical core issues, such security-based problems cannot exclude. Concerning their tremendous parallel computational capability, encryption and decryption algorithms challenging to put in force using digital computer systems may be feasible using DNA ones. Researchers will inherit computational stability in the DNA system if these schemes can survive quantum computer attacks. Therefore, DNA cryptography cannot wholly exclude traditional cryptography and can build a hybrid cryptosystem from it.

Defense Specifications

Researchers have the same cryptographic qualities regardless of the various types of DNA and classical crypto. The sender and recipients of the DNA-encrypted version of the discussion are often two organizations that have acquired secret keys safely or authenticated and then connect securely with each other across risky or unauthenticated routes. The necessity for security must also be based on Kerchoff's notion that one should find a guarantee on the decryption key's secrecy. Except for the decryption key, the attacker should remove all encrypted and decrypted data. After no attacker can crack it, the cryptographic system can be secure (Diffie & Hellman, 1976). It must be assumed, in particular, that the attacker is familiar with the designer's basic biological methods and possesses sufficient knowledge and excellent laboratory equipment to replicate the designer's operations. The only thing the intruder doesn't know is the key. Certain biological materials or preparation processes are often part of the DNA code system, sometimes even under experimental conditions.

Goal of Current Research

The current research target for DNA cryptography must often be protection and feasibility, and secondly, storage density. A cryptosystem of sound must be comfortable and smooth to put in force. Developing present-day biological technology enables DNA statistics to be articulated, even though the applicable work is still at its introductory level. The nanoscopic DNA is usually challenging to work at once. Scientists can perform DNA effortlessly with the help of limited enzyme sort's best after amplification of DNA strands with amplification technology, including PCR. It's also hard to store all the worldwide information with the brand new technologies to aid in using more than one gram of DNA. Until the most uncomplicated requirement is to increase storage capability, it's challenging to incorporate DNA cryptography at the present-day degree of method. Utilizing the colony assets of considerable DNA for cryptographers is extra realistic. For example, storing DNA chip data and hybridization statistics makes I/O operations faster and more convenient. Performing this method is more accessible than encoding the message as nucleotides, but the storage density is much lower.

Key Challenges

Meanwhile, the main task for DNA cryptographers is to lay the fundamental foundations and gain practical knowledge. Researchers may prove the massive parallelism, incredible strength efficiency and verify the notable density of records found in DNA. It prompts a rethinking of computers and cryptography in the context of DNA. The modern objective or dilemma is to realize and maximize one's potential fully. However, related research is still in the early stages. No reliable theory has been developed for every DNA and cryptographic calculation. Modern biology emphasizes experiment rather than theory. Based on this issue, there may not be an appropriate way to assess the toughness of biological concerns and the accompanying cryptographic system's security level. Like computational complexity, finding such a method is imminent. The most important aspect of the intervention time is finding an appropriate DNA location for calculation and encryption. To lay a theoretical foundation and accumulate experience, it is possible to build a realistic and straightforward DNA cryptography system under this premise.

COMPARATIVE ANALYSIS OF THIS STUDY

According to the research on various methodologies, there are several parameters that differ between each methodology is followed in Table 2.

Finally, the delicate balance between security and computing must be maintained. The technique must be capable of decreasing calculations in terms of computing time while still achieving the aim of maintaining security in a greater ratio. An efficient system is one that performs an average number

Table 2.
Comparison of DNA cryptosystem methodology

Authors	Functionalities	Advantages	Disadvantages	Operators
Aich et al, 2015	The XOR procedure is crucial in this method.	Because the OTP key is utilised, the attacker will not be able to access the key or the message.	In proportion to the size of the message, the computational complexity rises.	Arithmetic process
Cui et al., 2014a	Primers and extra complimentary sequences are used.	To ensure secure communication, less processing is required.	To get the original plaintext, the receiver must do PCR amplification and sequencing.	Biological process
Cui et al., 2014b	Primers and the Microdots method are used.	Physical concealment, stepwise transmission, mathematical security, and biological security are the four layers of protection provided.	Filter papers are required to implement encryption in real time.	Biological process
Jain & Bhatnagar, 2014a	Spiral Transposition is used, as well as a DNA sequence dictionary.	It's simple to set up and takes a variety of data input types.	Encryption does not rely heavily on biological activities.	Arithmetic process
Jain & Bhatnagar, 2014b	This article provides an overview of the different security algorithms.	Provides information on various security algorithm parameters.	The parameters that are compared are mostly concerned with the operators engaged.	Arithmetic and biological processes are both involved.
Jiang & Yin, 2013	DNA cryptography, Quantum cryptography, and classical encryption are discussed.	In comparison to quantum and classical encryption, the effectiveness of DNA cryptography is determined.	The implementation of the DNA cryptosystem is dependent on a biological laboratory. DNA passwords aren't re-usable.	Arithmetic and biological processes are both involved.
Misbahuddin & Mohammed, 2014	Splicing and indexing methods are employed, as well as complementary DNA sequence rules.	This technique entails a simple dynamic implementation of the algorithm.	The iterative process takes a long time.	Both Arithmetic and Biological process
Lu et al., 2007	Encryption and decryption keys are generated using DNA probes. It also uses a DNA chip to conduct a DNA hybridization method.	Forms a symmetric cryptosystem in its entirety.	Because it is entirely a biological process that necessitates biological resources, it is less adaptable to real-time applications.	Biological process
Yang et al., 2014	The strand displacement mechanism is included, as well as the detection of fluorescent signals.	Even though it involves the self-assembly structure of DNA, it obtains the correct encrypted text for the provided plain text.	Encryption and decryption are mostly accomplished using laboratory equipment.	Biological process

of calculations in a short amount of time. It would also be able to calculate in the least amount of time with the least amount of memory space for optimum speed and security that suits plaintext with uppercase and lowercase alphabets, digits, and special characters of whatever bigger size it might be. As a result, if a solution is developed that contains all of the aforementioned properties, it will be implemented in real time, resulting in a significant advancement in DNA cryptography. DNA cryptography is a new area of study that aims to improve data secrecy. Despite the fact that several DNA cryptosystems exist, finding one that meets all of the requirements is difficult. Another difficult aspect is establishing a set of measurements to assess the cryptosystem's complexity.

CONCLUSION

First of all, it is best to conduct DNA cryptography research, and several issues can resolved. However, DNA cryptography has distinct benefits over other cryptography methods because of its inherent broad parallelism, ultra-high energy efficiency, and large fact density. According to Adleman, "DNA computers like the one presented here illustrate that biological molecules (nucleic acids, proteins, etc.) can be used for recreational uses which are clearly non-biological. For these purposes, these molecules constitute an untapped legacy of 3 billion years of evolution, and their further discovery has great potential" (Ravinderjit et al., 2002). DNA binary strands show off capability for application in the world of cryptography. DNA-based encryption is a widespread step in opposition to the present-day techniques of encryption. However, DNA cryptography remains at its infancy degree and requires specific research with extensive discussion. Processes that had arisen on the molecular stage might be accomplished outdoor to the maximum. DNA-based algorithms for checking message integrity need to be established. Researchers should implement an appropriate programming language or module should be implemented mainly for DNA cryptography.

REFERENCES

Adithya, B., & Santhi, G. (2019, December). Bio-inspired Deoxyribonucleic Acid based data obnubilating using Enhanced Computational Algorithms. In *International Conference on Computing Networks, Big Data and IoT (ICCBI 2019)* (Vol. 49, pp.597-609), Springer.

Adithya, B., & Santhi, G. (2021). DNA Computing using Cryptographic and Steganographic Strategies. *Data Integrity and Quality*, 1-19. 10.5772/intechopen.97620

Adleman, L. M. (1994). Molecular computation of solutions to combinatorial problems. *Science*, *266*(5187), 1021–1023. doi:10.1126/science.7973651 PMID:7973651

Adleman, L. M., Rothemund, P. W., Roweiss, S., & Winfree, E. (1999). On applying molecular computation to the Data Encryption Standard. *Journal of Computational Biology*, *6*(1), 53–63. doi:10.1089/cmb.1999.6.53 PMID:10223664

Aich, Sen, Dash, & Dehuri. (2015). A Symmetric Key Cryptosystem Using DNA Sequence with OTP Key. *Advances in Intelligent Systems and Computing*, 207 – 215.

Bancroft, C., Bowler, T., Bloom, B., & Clelland, C. T. (2001). Long-Term storage of information in DNA. *Science*, *293*(5536), 1763–1765. doi:10.1126/science.293.5536.1763c PMID:11556362

Bennett, C. H. (1992). Quantum cryptography using any two nonothogonal states. *Physical Review Letters*, *68*(21), 3121–3124. doi:10.1103/PhysRevLett.68.3121 PMID:10045619

Bennett, C. H., & Brassard, G. (1984). Quantum cryptography: Public-key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (pp. 175-179). Bangalore Press.

Bennett, C. H., Brassard, G., & Ekert, A. K. (1992). Quantum cryptography. *Scientific American*, *267*(4), 50–57. doi:10.1038/scientificamerican1092-50

Celland, C. T., Risca, V., & Bancroft, C. (1999). Hiding messages in DNA microdots. *Nature*, *399*(6736), 533–534. doi:10.1038/21092 PMID:10376592

Cui, G., Han, D., & Wang, Y. (2014a). *An Improved Method of DNA Information Encryption. BIC-TA, CCIS*. Springer.

Cui, G., Han, D., & Wang, Y. (2014b). *An Encryption Scheme Based on DNA Microdots Technology. BIC-TA, CCIS*. Springer.

Daemen, J., & Rijmen, V. (2002). *The Design of Rijndael: AES the Advanced Encryption Stand*. Springer-Verlag. doi:10.1007/978-3-662-04722-4

Debao, L., & Ping, X. (1994). *Theory and Methods of Recombinant DNA*. Zhejiang Science and Technology Publishing Co.

Diffie, W., & Hellman, M. (1976). New directions in cryptography. *IEEE Transactions on Information Theory*, *22*(6), 644–654. doi:10.1109/TIT.1976.1055638

Ekert, A. K. (1991). Quantum cryptography based on Bell's theorem. *Physical Review Letters*, *67*(6), 661–663. doi:10.1103/PhysRevLett.67.661 PMID:10044956

Elgamal, T.EIGamal. (1985). A public-key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, *31*(4), 469–472. doi:10.1109/TIT.1985.1057074

Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., & Solas, D. (1991). Light-directed, spatially addressable parallel chemical synthesis. *Science*, *251*(4995), 767–773. doi:10.1126/science.1990438 PMID:1990438

Gehani, A., LaBean, T. H., & Reif, J. H. (2000). DNA-based cryptography. *Dimacs Series In Discrete Mathematics & Theoretical Computer Science*, *54*, 233–249. doi:10.1090/dimacs/054/19

Gibson, D. G. (2011). Enzymatic assembly of overlapping DNA fragments. *Methods in Enzymology*, *498*, 349–361. doi:10.1016/B978-0-12-385120-8.00015-2 PMID:21601685

Gifford, D. K. (1994). On the path to computation with DNA. *Science*, *266*(5187), 993–994. doi:10.1126/science.7973681 PMID:7973681

Grovel, L. K. (1996). Quantum mechanics algorithm for database search. In *Proceedings of the 28th ACM Symposium on the Theory of Computation* (pp. 212-219), New York: ACM Press.

Guarnieri, F., Fliss, M., & Bancroft, C. (1996). Making DNA add. *Science*, *273*(5272), 220–223. doi:10.1126/science.273.5272.220 PMID:8662501

Guozhen, X. I. A. O., Mingxin, L. U., Lei, Q. I. N., & Xuejia, L. A. I. (2006). New field of cryptography: DNA cryptography. *Chinese Science Bulletin*, *51*(12), 1413–1420.

Jain & Bhatnagar. (2014a). A Novel DNA Sequence Dictionary method for Securing Data in DNA using Spiral Approach and Framework of DNA Cryptography. *IEEE, ICAETR*.

Jain & Bhatnagar. (2014b). Analogy of Various DNA Based Security Algorithms Using Cryptography and Steganography. *IEEE, ICICT*.

Jiang & Yin. (2013). The Advantages and Disadvantages of DNA Password in the Contrast to the Traditional Cryptography and Quantum Cryptography. *Bio-Inspired Computing: Theories and Applications*, 307-316.

Kari, L. (1997). DNA Computing: Arrival of Biological Mathematics. *The Mathematical Intelligencer*, *19*(2), 9–22. doi:10.1007/BF03024425

Leier, A., Richter, C., Banzhaf, W., & Rauhe, H. (2000). Cryptography with DNA binary strands. *Bio Systems*, *57*(1), 13–22. doi:10.1016/S0303-2647(00)00083-6 PMID:10963862

Lipton, R. J. (1995). Using DNA to solve NP-complete problems. *Science*, *268*, 542–545. doi:10.1126/science.7725098 PMID:7725098

Liu, Q., Wang, L., Frutos, A. G., Condon, A. E., Corn, R. M., & Smith, L. M. (2000). DNA computing on surfaces. *Nature*, *403*(6766), 175–179. doi:10.1038/35003155 PMID:10646598

Lodish, B. A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, S. L., & Darnell, J. (2003). Molecular Cell Biology (5th ed.). W. H. Freeman and Co.

Lu, Lai, Xiao, & Qin. (2007). Symmetric-key cryptosystem with DNA technology. *Information Sciences. Science in China*, 324–333.

Misbahuddin, M., & Mohammed, H. N. (2014). *DNA for Information Security: A Survey on DNA Computing and a Pseudo DNA Method Based On Central Dogma of Molecular Biology. In ICCCT*. IEEE.

National Institute of Standards and Technology. NIST FIPS PUB 46-2. (1993). Data Encryption Standards. U.S. Department of Commerce.

Ouyang, Q., Kaplan, P. D., Liu, S., & Libchaber, A. (1997). DNA solution of the maximal clique problem. *Science*, *278*(5337), 446–449. doi:10.1126/science.278.5337.446 PMID:9334300

Paun, R. G., & Salomaa, A. (1998). DNA Computing: New Computing Paradigms. Berlin: Springer.

Pease, A. C., Solas, D., Sullivan, E. J., Cronin, M. T., Holmes, C. P., & Fodor, S. P. (1994). Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(11), 5022–5026. doi:10.1073/pnas.91.11.5022 PMID:8197176

Ravinderjit, S., Braich, R., Chelyapov, N., Johnson, C., Rothemund, P. W. K., & Adleman, L. M. (2002). Solution of a 20-Variable 3-SAT problem on a DNA Computer. *Science*, *266*, 499–502. PMID:11896237

Rivest, R. L., Shamir, A., & Adleman, L. M. (1978). A method for obtaining digital signatures and public-key. *Cryptosystems Communications of the ACM*, *21*(2), 120–126. doi:10.1145/359340.359342

Roweis, S., Winfree, E., Burgoyne, R., Chelyapov, N., Goodman, M. F., Rothemund, P. W. K., & Adleman, L. M. (1998). A sticker based model for DNA computation. *Journal of Computational Biology*, *5*(4), 615–629. doi:10.1089/cmb.1998.5.615 PMID:10072080

Sakamoto, K., Gouzu, H., Komiya, K., Kiga, D., Yokoyama, S., Yokomori, T., & Hagiya, M. (2000). Molecular computation by DNA hairpin formation. *Science*, *288*(5469), 1223–1226. doi:10.1126/science.288.5469.1223 PMID:10817993

Schena, M., Shalon, D., Davis, R. W., & Brown, O. P. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, *270*(5235), 467–470. doi:10.1126/science.270.5235.467 PMID:7569999

Seeman, N. C. (2004). Nanotechnology and the double helix. *Scientific American*, *290*(6), 34–43. doi:10.1038/scientificamerican0604-64 PMID:15195395

Shalon, D., Smith, S. J., & Brown, P. O. (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Research*, *6*(7), 639–645. doi:10.1101/gr.6.7.639 PMID:8796352

Shannon, C. E. (1949). Communication theory of secret systems. *The Bell System Technical Journal*, *28*(4), 656–349. doi:10.1002/j.1538-7305.1949.tb00928.x

Shor, P. W. (1994). Algorithms for quantum computation: Discrete log and factoring. In *Proceedings of the 35th Symposium on Foundations of Computer Science* (pp. 124-134), Los Alamitos, CA: IEEE Computer Society Press. doi:10.1109/SFCS.1994.365700

Simon, D. (1994). On the power of quantum computation. In *Proceedings of the 35th Symposium on Foundations of Computer Science* (pp. 116-123), Los Alamitos, CA: IEEE Computer Society Press. doi:10.1109/SFCS.1994.365701

Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., & Weiner, A. M. (1987). Molecular Biology of the Gene (4th ed.). The Benjamin/Cummings Publishing Co., Inc.

Wiesner, S. (1983). Conjugate coding. *SIGACT News*, *15*(1), 78–88. doi:10.1145/1008908.1008920

Yang, J., Ma, J., Liu, S., & Zhang, C. (2014). A molecular cryptography model based on structures of DNA self-assembly. *Chinese Science Bulletin*, *59*(11), 1192–1198. doi:10.1007/s11434-014-0170-4

Santhi G. is currently working as an Assistant Professor in the Department of Information Technology under Puducherry Technological University. Area of Specialization is Computer Networks, Information Security, Wireless Networks, Computer Architecture.