A Method to Encrypt Information with DNA-Based Cryptography

Mohammadreza Najaftorkaman\textsuperscript{1}, Nazanin Sadat Kazazi\textsuperscript{2}

\textsuperscript{1}School of Information and Communication Technology, Griffith University  
Gold Coast, Queensland, 4222, Australia  
\textsuperscript{2}Advanced Informatics School (AIS), UNIVERSITI TECHNOLOGI MALAYSIA (UTM),  
Kuala Lumpur, Malaysia  
mohammadreza.najaftorkaman@griffithuni.edu.au, sknazan2@live.utm.my

ABSTRACT

DNA cryptography is a relatively new paradigm that has attracted great interest in the field of information security. Although there are many problems in DNA cryptography, scientists are trying to solve them because they believe that, with the extraordinary information density and the vast parallelism that are inherent in DNA computers, it is possible to make a secure system. In this study, we primarily focus on modern cryptography, which is based on quantum and DNA cryptography. Furthermore, a novel method to encrypt data by using DNA-based cryptography is discussed. In our study, DNA coding technology is used to convert binary data to DNA strings. Finally, the proposed DNA cryptography algorithm’s strength is evaluated based on the properties of DNA strands and probability theories.

KEYWORDS

Information security, cryptography, nano-cryptography, DNA cryptography, DNA coding technology, DNA microchip

1 INTRODUCTION

Cryptography is the process used to encrypt plaintext into incoherent ciphertext [3-6]. In this process, the algorithm and key play vital roles. Information security experts have focused on nano-cryptography to create more secure systems. When scientists found that binary computers (digital computers) have various physical constraints, especially in data storage and computation processes, they concentrated on DNA computers (biomolecular computers) and quantum computers. DNA cryptography is a new science, of which deoxyribonucleic acid (DNA) is the most important feature [7, 8]. Cryptographers have worked on DNA cryptography to solve common limitations to make a system which is immune to popular attacks like brute-force attacks, dictionary attacks, etc [9].

The main achievement of this study is identifying the DNA cryptosystem, which is new science in information security. Moreover, we have designed and implemented a new DNA cryptography algorithm and elaborated its security compared to other current DNA cryptography algorithms. In this paper, we used DNA cryptosystem concepts based on the classic Vigenere cipher. Below, section 2 presents related research studies in DNA cryptography and explain research background. The five steps to encrypt and decrypt binary information based on DNA cryptography are clarified in section 3. In section 4, we focus on the strength of the proposed cryptosystem based on the properties of DNA technology and probability theory. Conclusions are presented in section 5.
2 RELATED WORKS

Nano-cryptography is divided into quantum and DNA cryptography. The first quantum cryptography issue was presented by Bennett and Brassard in 1984 [1]. The fundamental concept of quantum cryptography is sending secret key in the form of photons through an insecure channel. Binary data (zero and one) is encoded to a quantum state based on physics theories. Bennett and Brassard proposed the first quantum cryptography protocol, which is known as BB84, to exchange secret key. Quantum cryptography uses quantum bits (Qbit) to show the data. Actually, a Qbit can be 0 and 1 at the same time. The general theory of the Qbit is illustrated in Figure 1 [2].

Quantum cryptography is also well known as quantum key distribution (QKD). There are two channels in this system. The first channel is used to transmit the quantum secret key with a single photon. The second channel is a public channel like a telephone line or the Internet used to exchange cryptography protocols. Figure 2 shows the quantum key distribution [1].

![Figure 1. Quantum bit (Qbit) [2]](image)

The second area of nano-cryptography is DNA cryptography. The background of this new cryptography technique is DNA computing. DNA computing is a novel computational model which is based on some vital chemical reaction. In 1994, Adleman introduced a novel approach in the information era [10]. His first essential experiment was solving the directed Hamilton path problem in the molecular biology lab. This experiment showed the possibility of doing computations at a molecular level. He tried to solve one of the popular NP-complete problems, the directed Hamilton path problem, based on a DNA algorithm. Based on Adleman’s experiment, many scientists have tried to solve NP-problems with DNA computers [11-13].

In the first experiments on DNA computing, scientists focused on NP-problems. They tried to find an efficient algorithm based on DNA computing concepts to solve NP-problems like the directed Hamilton path, SAT problem, N-Queen problem, etc. Boneh [14] was the first person to present the application of a DNA algorithm in cryptography. He used DNA computing to break Data Encryption Standards (DES). After Boneh, further research was done by different scientists. DNA cryptography methods based on a one-time-pad cryptosystem were shown by Gehani [15], who used a one-time pad and the properties of natural DNA to encrypt information.

![Figure 2. Quantum key distribution [1]](image)

He created a DNA string that is divided into a plaintext block and ciphertext block with a stop codon between the ciphertext and plaintext. In the encoding process, DNA plaintext is mixed into the corresponding complementary sequence[16].

Clelland [16] used a substitution cipher to encode characters to DNA strands. The deciphering process is reversible. In fact, Clelland used a table that converts letters of the English alphabet to DNA sequences. For instance, letter
‘A’ converts to codon ‘CGA’, letter ‘B’ converts to ‘CCA’, etc. Leier [17] tried to encode binary information into DNA strands. A short DNA string represents 1, while another based on the properties of that sub-string represents 0. Two very short DNA sequences are used as the start and end points. At the end of encryption, the DNA sequence is mixed by dummy strands. If we have the primary sequence we can decrypt it. Wong [18] presented an encryption algorithm based on DNA steganography. He used living organisms to store data. One of the disadvantages of Wong’s experiments is using a natural organism, which leads to some errors in the experiment because it is related to certain chemical elements and temperature.

Heider [19] proposed an algorithm to encode binary information into a DNA sequence. He used DNA coding technology that converted 00 to base ‘T’, 01 to ‘G’, 10 to ‘C’ and 11 converted to a DNA-based ‘A’. After that, the result of this process was mixed by a DNA-based sequence. The advantage of Heider’s algorithm is the ability to make use of a binary encryption algorithm like AES. Kang Ning [20] proposed a pseudo DNA cryptography method that is based on the central dogma of molecular biology. Although this project needs improvement, it presents a new view of DNA cryptography.

The most popular DNA cryptosystem uses the sticker model, which is based on Watson-Crick complementarity. This model was proposed by Roweis in 1996 [21, 22]. The sticker model has two major advantages. It does not use enzymes and operations do not need strand extension. Several molecular sticker models have been studied [23, 24]. Chen [25] proposed a DNA algorithm to break DES based on the modified sticker model. The basic operations like XOR and bit substitution in this algorithm are described in [26]. Chen claimed that the probability of success of his algorithm is 100% and its computational complexity is \( (n^2) \), so this algorithm is better than all the other algorithms [14].

DNA cryptosystems have been studied by many scientists [27-30]. Among these studies some experiments have important advantages which are based on the use of DNA probes and DNA chips. The first experiment is the symmetric-key cryptosystem with DNA technology [20]. The second is the DNA asymmetric-key cryptosystem (DNA-PKC) [26]. Next, Lu [31] proposed a DNA symmetric-key cryptosystem (DNASC). He used DNA chips (microarray) and DNA probes in his experiment. He claimed that DNASC is not similar to classical DNA technology, because of some problems with the ability of the classical model of DNA computation to react in a parallel way at the same time. A large amount of DNA probes are used to hybridize in parallel steps. Although there are many DNA cryptosystems based on the DNA coding sequence to achieve their purposes, DNASC used DNA chip technology. This novel technology offers many advantages in detecting, reading and extracting DNA data [31]. Lai [26] extended Lu’s idea and proposed a DNA asymmetric-key cryptosystem (DNA-PKC) based on novel approaches which used DNA probes and DNA chips (microarray). He used Lu’s experiment to select probes and a hybridized process. DNA-PKC involves some steps based on public-key cryptography.

DNA coding technology is another concept in cryptography that is intended to encode binary data to a DNA strand and vice versa. There are different ways to encode binary data [24, 32, 33]. Binary data can be encoded in DNA by using Oligonucleotide sequences or the alphabet. By using this approach, it is easy to implement DNA cryptography in binary computers. The result of DNA digital coding can be adapted to existing binary computers. Also in this way, conventional cryptography systems like RSA and DES are used and implemented in DNA cryptography concepts [15, 34]. For instance, one of the simple coding rules is described below. It is known that DNA sequences contain the four basic letters A, C, G, and T. This method has been used in various studies [35, 36].

- 00 converts to A
- 01 converts to C
- 10 converts to G
- 11 converts to T
3 THE PROPOSED METHOD

There are five main steps in implementing the proposed DNA cryptography: data pre-processing, key generation, encryption, decryption, and data post-processing.

3.1 Data pre-processing

In this step binary data is converted to a DNA string. There are many approaches in this field to convert binary data to a DNA string and these are known as DNA coding technology. We propose a new DNA coding technology to convert binary data to DNA strings. The DNA coding rule is shown in Table 1.

Table 1. DNA coding technology

<table>
<thead>
<tr>
<th>Binary Data</th>
<th>DNA Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>AA</td>
</tr>
<tr>
<td>01</td>
<td>T</td>
</tr>
<tr>
<td>10</td>
<td>C</td>
</tr>
<tr>
<td>11</td>
<td>GG</td>
</tr>
<tr>
<td>0</td>
<td>A</td>
</tr>
<tr>
<td>1</td>
<td>G</td>
</tr>
</tbody>
</table>

The concept of the table above is as follows: if we have ‘00’ in the binary string it is converted to ‘A’, or if we have ‘01’ it is converted to ‘T’. This process continues to convert all the binary information into a DNA sequence. If the length of the binary data is not even, the last two rows help us to convert correctly. For example, if the value of binary string is:

001011010010001

The value of this binary string, which is based on the proposed DNA coding technology, is:

AATAACCTCTCCCTCTAACTAACCTAACCTAAC

3.2 Key generation process

In this step, we used the NCBI bank (National Center for Biotechnology Information), which is the master bank of the human genome. We can therefore search many different kinds of DNA string. NCBI provides a sample database of DNA strings and we used MATLAB software to extract from the NCBI database. There is a function ‘getbank’ to load the DNA string from the NCBI database. The following function extracts the DNA sequence from the NCBI bank.

```matlab
Mitochondria=
getgenbank('NC_001807','SequenceOnly',true);
```

After that, we used variable mitochondria to produce our secret key. One random number was generated by MATLAB software that shows the start index of the DNA strand. Next, all the DNA nucleotides were extracted from the start point. In our study we chose a DNA secret key with a key length of 256. For example, the output of the getgenbank function is:

TCTATCACCCTATTAAACCACCTCACGGGAGCTCCTCATGCTTTGGGATTCCTGCTGCCGAGGTTGGAGCCCGGAGCAACGTGTTAAGTGTGTTAATTATAATTGAATGGTACATCTATTACACCCTATTTGTCGAGAAGGCTGAGGCGAGGACACCTCTAGTGACATTTTCGTCTTGGGGCTGAGGACGACGAGCATTTTGGGATTCCTGCTGCCGAGGTTGGAGCCCGGAGCAACGTGTTAAGTGTGTTAATTATAATTGAATGGTACATCTATTACACCCTATTTGTCGAGAAGGCTGAGGCGAGGACACCTCTAGTGACATTTTCGTCTTG

We used the selected DNA sequence as follows:

TCTATCACCCTATTAAACCACCTCACGGGAGCTCCTCATGCTTTGGGATTCCTGCTGCCGAGGTTGGAGCCCGGAGCAACGTGTTAAGTGTGTTAATTATAATTGAATGGTACATCTATTACACCCTATTTGTCGAGAAGGCTGAGGCGAGGACACCTCTAGTGACATTTTCGTCTTG

This is a short sequence of the DNA string. It is used to generate the key that is used in encryption and decryption. The key is used to encrypt the binary string and the encrypted string is a DNA sequence. The DNA sequence is then used to decrypt the binary string and the decrypted string is the original binary string.
GTCAATATTACGGCGAACATACCTACTA
AAGTGTTAATTAAATTAA

We used the ‘Rand’ function, a standard MATLAB function, to identify the start index of the DNA secret key.

\[
\text{startIndex} = \text{fix}(10 \times \text{rand});
\]

After that we added 255 to the “startIndex” variable to identify the end index of the DNA strand. Therefore, the DNA secret key is:

GTCTATACCCATTAACCACCTACGGGAG
CTCTCCATGACATTGTATTTTCGTCTGGG
GGGTGTCACACGAGATCATGGCGAGACG
CTGGAGCCCGAGCACACCTATGTCGCAAGTA
TCTGTCTTTGATTTCTGCCCTTACTATTAT
TTATCGACACCTACGTCAATATTACAGGCG
AACATACCTACTAAAGTGTATAATTAAATT
AATGCTTGTAGGACATAATAATAAAAATG
GAATGTCTGCACACGCCGC

3.3 Encryption process

In this step, the algorithm was developed to encrypt DNA strand plaintext according to our secret key. The output of this process is a DNA strand ciphertext. The encryption process was based on the Vigenere cipher, which is a polyalphabetic cipher [36]. The DNA strand plaintext was processed based on the secret key to make a DNA strand ciphertext. First, we produced a DNA-Vigenere table based on the properties of the Vigenere cipher [37], as shown in Table 2.

<table>
<thead>
<tr>
<th>Table 2. DNA-Vigenere table</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>G</td>
</tr>
</tbody>
</table>

The encryption algorithm has two main features. The first feature is plaintext and the second is the secret key. The characters of the plaintext show the column number of the DNA-Vigenere table and the character of the secret key shows the row number of the table after replacing the character in the Vigenere table with real plaintext. For example, if the first character of the plaintext is ‘A’ and the first character of the secret key is ‘G’ that means column 1, row 4. Therefore, we replace the first character of plaintext (A) with the V value of (4,1), that is, G.

3.4 Decryption process

The input of this process is DNA strand ciphertext and the output is DNA plaintext. In this step the secret key was used to decrypt the ciphertext. In fact, the function of the decryption process is the reverse process of decryption. The DNA-Vigenere table was used to convert the cipher strand to plaintext by the secret key.

3.5 Data post-processing

We obtained the DNA sequence plaintext from the decryption process. Now, data post-processing converts the DNA data to binary plaintext. This process is based on the same DNA coding technology that is used in the data pre-processing step.

4 DISCUSSION

The proposed algorithm is based on the classic Vigenere cipher. In fact, we mapped the main rules of the Vigenere cipher on the DNA cryptography area and made the new algorithm. We tried to alleviate the weaknesses of the classic Vigenere cipher and add more security features to it. The input of this cryptosystem is binary plaintext. The data pre-processing step converts the binary data to DNA plaintext. After that, the DNA plaintext is encrypted by the encryption function with our DNA secret key. The decryption process is the reverse function of encryption.
4.1 Security evaluation

Nowadays, nanotechnology is bringing a new revolution in science. Since scientists found that binary computers have many physical limitations, especially in data storage and computation, they have concentrated on DNA computers (biomolecular computers) and quantum computers and tried to implement this new science in the information security field. They claimed that by using the two major properties of these computers, namely their extraordinary information density and vast parallelism, they can create good, secure systems [10]. DNA cryptography is a new concept that needs many improvements. Although there are still problems with DNA cryptography, many scientists are trying to solve them because they believe that, with the characteristics of DNA computers, they have more advantages than conventional cryptography.

The cryptosystem of this project is based on the Vigenere cipher, which is a polyalphabetic cipher. Polyalphabetic ciphers are multi-substitution ciphers, which means that each letter in the plaintext is substituted in different forms. However, the Vigenere cipher has some problems in classic systems. If we use Vigenere cipher concepts within the DNA concepts we can reduce the weaknesses and the previous vulnerability.

The first problem of the Vigenere cipher is its vulnerability to frequency analysis. The Vigenere cipher uses the English alphabet so it is obvious that with frequency analysis we can guess the correct letter of the ciphertext. For example, the letter in the ciphertext that occurs most frequently should be letter ‘E’, which is the most frequent letter in the English language. This approach can be continued to decrypt all of the ciphertext.

In the proposed DNA cryptosystem, the weakness against frequency analysis is solved. In this algorithm first the binary data is converted to DNA strings by the DNA coding technology. For example, a binary data string like ‘01001011’ is converted to ‘TAACGG’. In the second step of encryption we are working on a DNA string which is made from the four main letters: ‘A’, ‘C’, ‘G’, ‘T’. Therefore, we cannot guess the frequency of letters as in the previous cipher. For instance, in previous approaches the most common letter in the ciphertext is denoted as ‘E’ in the English language, but now we cannot find any relationship between letters. This means that, if we have a cipher such as:

GGGCTAGCTGGATAGCTAGATTAAAGT
AATAGTAGCCCTTGACCCCCGGCGCTTAAC
GTAAAGGTGTGACAGTATATAGGACTGAG
ACGTCATGC

We can see that the letter ‘G’ is the most common in the ciphertext. However, we cannot find any relationship with the letters together to guess the correct plaintext (see Figure 3).

![Figure 3. Letter frequency](image)

Another famous problem with the Vigenere cipher is the repetition of its secret key where the key length is short. If the length of the Vigenere key is guessed, all the ciphertext is breakable because the length of each section is defined and after that each section is used to extract the correct plaintext by using the frequency analysis approach. For this weakness, the main test to find the correct key length is the Friedman test or Friedman method.

The Friedman test uses an index of coincidence which calculates the evenness of cipher letters to break the main ciphertext. The purpose of the Friedman test is to find the key length.

\[
\text{Key length} = \frac{K_p - K_f}{K_0 - K_f}
\]
\[ K_p = \text{The probability that two randomly chosen ciphertext elements are the same. In English, this value is 0.067.} \]

\[ K_r = \text{Probability of coincidence for random selection from the alphabet: in DNA coding technology we have the four letters A, C, T, and G. Therefore the value is } \frac{1}{4}. \]

\[ K_0 = \text{Coincidence rate:} \]

\[ \frac{\sum_{i=1}^{c} (f_i \times f_{i-1})}{N(N-1)} \]

\'c\' is size of alphabet, \'N\' is length of ciphertext, \( f_i \) is letter frequency. We explain one example that shows how to secure the cipher against the Friedman test.

**Binary plaintext=244:**

01010100111010000110101110011001000000
0110100101111001101000000110101111001
00100000100100111100011000001101000000
0110001101111100110011100001111001000
0010000011000011011001100011011110001
0111001001110110011010110011101011001
0011000010110100110011011011011101

**DNA string plaintext=173:**

TTTATCCAACTCTTGGGAGGAAACAAAT
CCTTGGGAAAGGAAACAAAATCGTGTTGCTAA
CAAAATAATAAATAAGCTAAATAAACAA
ATCAAGGTGGAACTGACCTTGGGAATATGG
AAAAACAAATCAATCAGAATCGTCTCG
GGTGGAACCTCTTGGTAATCCCAATCGGT

**Secret key=256:**

GGTCTATCCACCTATTAACCACCTCACGGGA
GCTCTCCATGCATTGATTTTCGCTCG
GGGTGTCGACGCGATAGCTTGGCGAGAC
GCTGGAGCCGAGGACACCCTATGTCGACGT
ATCTGCTTTTGGATTCTGCTACCTATTATT
TTATCGACACTACGTCTCAGATATTACGCG
GAACATCACCCTAATAGGTTGTTTAATTAT
TAATGCTTTGAGACATAAAATAAAATATT
GAATGCTCGACAGCC

**DNA string cipher text=173:**

AACGTGAAGCAGTCAGAGTGCTAAGGGA
CTACAGGAGGCTTTGAGGACATTCTGT
GGGGATGCCGTTGGTGCTCAGAGGAG

\[ \text{Index of coincidence:} \]

\[ \sum_{i=1}^{c}(f_i \times f_{i-1}) \]

\[ N(N-1) \]

\[ A (34*33) + C (36*35) + G (55*54) + T (48*47) = 1122+1260+2970+2256=7608 \]

\[ N (N-1) = 173*172=29756 \]

\[ IC=7608/29756=0.25567 \]

\[ \text{Key length} = \frac{K_p - K_r}{K_0 - K_r} = \frac{K_p - 0.25}{0.25567 - 0.25} \approx \infty \]

As we can see, the value of the index of coincidence is near to 0.25 so, based on the Friedman formula, the denominator is around 0, and the key length is \( \infty \). Therefore we cannot find the exact key length.

In the last part of this subsection, we imagine that the attacker is having difficulty finding the key length. Therefore, he tries to use the frequency analysis approach to find the correct plaintext. However, we have explained that with frequency analysis we cannot find any relationship between letters. Consequently, it will be very hard to find the plaintext with the secret key.

**4.2 Double layer security**

We can add another security layer for this cryptosystem. If we convert the output of the encryption algorithm (DNA string) into real molecular DNA and write it on a DNA microchip, we have a double layer of security. The first layer is computational security and the second is biological security. The DNA chip consists of a set of probes which can contain DNA sequences. Many scientists are working on DNA microchips and the DNA strands on these chips [26, 31, 38, 39]. Assume that we use one of the standard DNA microchips, which have 190 million DNA sequences. We try to embed the DNA key in the DNA microchip. The probability of finding the
correct DNA sequence among 190 million DNA sequences is calculated as follows:

\[
\text{Probability of find the correct key} = \frac{1}{1.90 \times 10^8}
\]

Therefore, it is very difficult to find the correct DNA sequence, and thus also to guess the correct secret key. Furthermore, extracting the correct sequence from the DNA chip is based on the correct chemical conditions which were used originally to write in the DNA chip. For example, if we want to embed the DNA sequence in a part of the DNA chip we have to follow certain chemical rules like using the appropriate temperature. If we do not know the correct temperature during embedding, the DNA strings will be separated and cannot survive. Consequently, the only way to extract the correct DNA sequence is by using the correct biological situation. As we can see, the attacker faces a double layer of security when looking for the correct key and then guessing the correct plaintext. Brute-force attack, which reads each spot of the DNA chip directly to collect the true sequence, is not able to find the correct DNA sequence because there are so many mixed molecules with different physical and chemical conditions like temperature or chemical materials.

Finally, we compare some popular DNA cryptography algorithms with the algorithm proposed in this study. Clelland [16], Gehani [15] and Wong [18] worked on natural DNA molecules but they could not convert binary data to a DNA strand (see Table 3). Therefore they could not use a binary encryption algorithm. As a result of working on natural DNA, their experiments had many errors. Heider [19] proposed an algorithm that used DNA coding technology and could also convert binary data to DNA strands. The properties of the algorithm proposed in this project are as follows:

- It is possible to convert binary data to a DNA strand (DNA coding technology). This operation was done in the first step of this study.

The fundamental algorithm of this research was the Vigenere cipher. First we converted binary data to a DNA strand and then we applied the Vigenere cipher. It is obvious that we can use DNA coding technology to apply binary algorithms to encrypt data. It is feasible to apply a binary encryption algorithm like DES in DNA cryptography.

### Table 3. Comparing of some famous DNA cryptography algorithms with the proposed algorithm

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Binary</th>
<th>Encryption</th>
<th>DNA chip</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clelland [16]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Gehani [15]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Wong [18]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Heider [19]</td>
<td>YES</td>
<td>YES</td>
<td>No</td>
<td>Medium</td>
</tr>
<tr>
<td>Proposed DNA-cryptography algorithm</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>Low</td>
</tr>
</tbody>
</table>

**Binary**: Binary information can be encoded  
**Encryption**: The use of a binary encryption algorithm like AES  
**Convert to DNA chip**: Using a DNA Chip  
**Error**: Difference between the output of decryption process and the original plaintext

It is possible to write the output of the encryption algorithm on a DNA microchip. Therefore, the amount of mutation errors is significantly reduced. After obtaining the result of the ciphertext, it is feasible to write the secret key and ciphertext on the DNA microchip. Based on working on DNA microchips and DNA coding technology, the amount of errors is reduced. These kinds of errors are related to biological concepts such as mutation.

### 5. CONCLUSIONS

The main achievement of this study is identifying a DNA cryptosystem, which is new science in information security. Furthermore, we
designed and implemented a DNA cryptography algorithm and elaborated its security compared to other current DNA cryptography algorithms. In this paper we used DNA cryptosystem concepts based on the Vigenere cipher. DNA cryptography aims to improve the weaknesses of current cryptography. Furthermore, we found that the DNA cryptosystem is more secure because it has two layers of security, which are computational and biological security. However, these concepts are new and they still need some improvement. Additionally, if we want to use these systems, we need facilities like DNA chips, and robots to prepare experimental set-ups.

The main contributions of this study are as follows:

- It is possible to convert binary data into a DNA strand (DNA coding technology).
- We can apply any binary cryptography algorithm in the DNA cryptography area.
- It is feasible to write a DNA string which is the output of our encryption algorithm to the DNA microchip. By using the DNA microchip, the number of biological errors decreases significantly.

For future work on DNA cryptography, further work on DNA microchips is suggested to provide a double layer of security. The limitation of this study is when converting the output of the encryption algorithm to real molecular DNA. To achieve this purpose, we need input from scientists from different areas, such as biologists, chemists, and computer scientists.

REFERENCES


