Multiple Sequence Alignment by Quantum Genetic Algorithm

Layeb Abdesslem, Meshoul Soham, Batouche Mohamed
LIRE laboratory, PRAI Group, University of Mentouri, Constantine, Algeria
layeb@yahoo.fr, meshoul@wissal.dz, batouche@wissal.dz

Abstract

In this paper we describe a new approach for the well known problem in bioinformatics: Multiple Sequence Alignment (MSA). MSA is fundamental task as it represents an essential platform to conduct other tasks in bioinformatics such as the construction of phylogenetic trees, the structural and functional prediction of new protein sequences. Our approach merges between the classical genetic algorithm and some principles of the quantum computing like interference, measure, superposition, etc. It differs from other genetic methods of the literature by using a small population size and a less iteration required to find good quality alignments thanks to the used quantum principles: state superposition, interference, quantum mutation and quantum crossover. Another attractive feature of this method is its ability to provide an extensible platform for evaluating different objective functions. Experiments on a wide range of data sets have shown the effectiveness of the proposed approach and its ability to achieve good quality solutions comparing to those given by other popular multiple alignment programs.

1. Introduction

Multiple Sequence Alignment (MSA) is basic task in bioinformatics. It consists in aligning several sequences in order to show the fundamental relationship and the common characteristics between a set of protein or nucleic sequences. This task is a fundamental platform for several other more complex tasks such as protein analysis, identification of functional sites in genomic sequences, structural and functional prediction of sequences and the construction of phylogenetic trees. Unfortunately, finding an accurate multiple alignment has been shown NP-hard [1]. Indeed, the MSA is an optimization problem which exhibits a great temporal and space complexity. Therefore several methods were proposed which can be grouped in three great classes [2]. The first class includes exact methods which use a generalization of Needleman algorithm [3] in order to align all the sequences simultaneously. Although exact methods give optimal solutions, their main shortcoming is their complexity which becomes even more critical with the increase of the number of sequences. The second class contains methods based on a progressive approach [4]. For these methods the multiple sequence alignment is built gradually according to a given order of the sequences, starting by the alignment of two sequences then it adds gradually the remaining sequences one by one to the preceding alignment. The progressive methods are simple, fast and generally give alignments of good qualities. However, their major disadvantage is the problem of the local minima and consequently they can lead to poor quality solutions. To overcome this problem, the iterative methods of the third class were showed to be promising. The basic idea is to start by an initial alignment and iteratively refines it through a series of suitable refinements called iterations. The process is reiterated until satisfaction of some criteria. Iterative methods can be deterministic or stochastic, depending on the strategy used to improve the alignment. The first stochastic iterative algorithm proposed in the literature uses an algorithm of simulated annealing [5]. However this algorithm is very slow and it is appropriate to be used as improver [2]. Later, several other iterative algorithms which use various strategies like Genetic Algorithms GAs [6], Tabu Search [7], were proposed. Concerning the deterministic iterative methods, they involve extracting the sequence one by one from multiple alignments and realigning them to the remaining sequences. The process is reiterated until it does not have more possible improvements. Although the iterative methods give generally more accurate alignments than the progressive methods, their major disadvantage is their high execution time.

One of the iterative methods that have been developed recently to solve this type of problem is Genetic Algorithms GA. It is a stochastic iterative algorithm which maintains a population of individuals. GA adapts nature optimizing principles like mechanics of natural selection and natural genetics. Each individual represents a feasible

1-4244-0054-6/06/$20.00 ©2006 IEEE

solution in the problem search space. Basically, a genetic algorithm consists of three essential operations: selection, crossover, and mutation. The selection evaluates the fitness of each individual and keeps the best ones among them. The others are removed from the current population. The crossover merges two individuals to provide new ones. The operator of mutation allows moving each solution to one of its neighbours in order to maintain a good diversity during the process of optimization. GA allows guided search that samples the search space. Although GAs have been showed to be appropriate for solving MSA problem [6], their computational cost seems to be a dissuasive factor for their use on large instances. To overcome this drawback and in order to get better speed and quality convergence, their implicit parallelism is exploited.

Quantum Computing QC is a new field in computer science which has induced intensive investigations and researches during the last decade on quantum mechanical computers and quantum algorithms. QC relies on the principles of quantum mechanics like qubit representation and superposition of states. QC is capable of processing huge numbers of quantum states simultaneously in parallel. QC brings new philosophy to optimization due to its underlying concepts. However, the quantum machines that the quantum algorithms require to be efficiently executed are not available yet. Recently, a growing theoretical and practical interest is devoted to researches on merging evolutionary computation and quantum computing [8, 9]. The aim is to get benefit from quantum computing potentials to enhance both efficiency and speed of classical evolutionary algorithms. One of the most successful quantum inspired algorithms is the Quantum inspired Genetic Algorithms QGA that have been proven to be better than conventional GAs.

The quantum evolutionary algorithm was first used in MSA problem in [10] to improve the quality of Clustal [11] alignments. The authors used only the mutation operators to create the diversification and the Weighted Sum of Pairs score as fitness criterion. In the present work, we propose a new algorithm for making alignment based on QGA. We have used in addition of mutation operators, the quantum Crossover operators in order to generate new alignments by combining two existing ones. By another hand, the Coffee function [12] was used to evaluate the quality of the potential alignments. We used a quantum representation to encode chromosomes. The advantage of this representation is the possibility to represent all possible alignment using only one chromosome. The other feature of the method is the application of quantum operators like measurement and interference. Our program can work on pure iterative method or by using a progressive\iterative method, the last are much suggested to get a good solution in small runtime. We have tested our approach on both the BALiBASE benchmark base [13] and RNA benchmarks BRALIBASE [14], and then we have compared our alignment results with other packages. The experimental results show that our approach can give a global alignment as good as better than the existent programs.

The remainder of the paper is organized as follows. In section 2, a formulation of the tackled problem is given. Section 3 presents some basic concepts of quantum computing. In section 4, the proposed method is described. Experimental results are discussed in section 5. Finally, conclusions and future work are drawn.

2. Problem Formulation

MSA can be formulated mathematically as follows: Given n sequences ($n \geq 2$) $S = \{s_1, s_2, ..., s_n\}$ defined over an alphabet Λ. The problem of MSA can be defined by specifying implicitly a pair (Ω, C) where Ω is the set of all possible solutions that is potentials alignments and C is a mapping $\Omega \rightarrow R$ called score of the alignment. Each potential alignment is viewed as a set $S' = \{s'_1, s'_2, ..., s'_n\}$ satisfying the following criteria:

- Each sequence s'_i is an extension of s_i and is defined over the alphabet $\Lambda' = \Lambda \cup \{-\}$. The symbol "–" denotes a gap. The deletion of gaps from s'_i leaves s_i.
- For all i, j: $length(s'_i) = length(s'_j)$.
- The score $C(S') = \sum_i \sum_j sim(s'_i, s'_j)$ is maximized where $sim(s'_i, s'_j)$ denotes some similarity between each pair of sequences s'_i and s'_j.

It is clearly that MSA is a combinatorial optimization problem. It appears to be impossible to obtain exact solutions in polynomial time. The main reason is that the required computation grows exponentially with the size of the problem. Therefore, it is often desirable to find near optimal solutions to these problems. Efficient heuristic algorithms offer a good alternative to accomplish this goal. Exploitation of the optimization philosophy and the parallel great ability of quantum computing is an attractive way to probe complex problems like MSA. Within this perspective, we are interested in applying quantum computing principles to solve MSA problem.

3. Quantum Computing And Overview

Quantum Computing is an interdisciplinary field which borrows techniques from many disciplines like physics, chemistry, computer science and mathematics. The QC

uses the specificities of quantum mechanics for the processing and the transformation of information. In 1980 Richard Feynman has shown the possibility of the use of quantum effects in the data processing. Later in 1994, Shor [15] demonstrated that QC can solve efficiently a NP-hard problem. Shor described a polynomial time quantum algorithm for factoring numbers. In 1996 Grover [16] has presented a quadratic algorithm for database search. Since that the QC has attracted widespread interest and has induced intensive investigations and researches since it appears more powerful than its classical counterpart. A particle according to principles of quantum mechanics can be in a superposition of states. By taking account of this idea, one can define a quantum bit or the qubit which can take value 0, 1 or a superposition of the two at the same time. Its state can be given by:

$$\Psi = \alpha\,|0\rangle + b|1\rangle \quad (3)$$

Where $|0\rangle$ and $|1\rangle$ represent the classical bit values 0 and 1 respectively; α and β are complex numbers such that

$$|\alpha|^2 + |b|^2 = 1 \quad (4)$$

The probability that the qubit collapses towards 1 (0) is $|\alpha|^2$ ($|b|^2$).This idea of superposition makes it possible to represent an exponential whole of states with a small number of qubits. According to the quantum laws like interference, the linearity of quantum operations and entanglement make the quantum computing more powerful than the classical machines. Each quantum operation will deal with all the states present within the superposition in parallel. For in-depth theoretical insights on quantum information theory, one can refer to [17].

In order to exploit effectively the power of quantum computing, it is necessary to create efficient quantum algorithms. A quantum algorithm consists in applying of a succession of quantum operations on quantum systems. Quantum operations are performed using quantum gates and quantum circuits. It should be noted that designing quantum algorithms is not easy at all. Yet, a powerful quantum machine is still under construction. By the time when a powerful quantum machine would be constructed, researches are conducted to get benefit from the quantum computing field. Since the late 1990s, merging quantum computation and genetic computation has been proven to be a productive issue when probing complex problems. Like any other GA, a Quantum Genetic Algorithm QGA relies on the representation of the individual, the evaluation function and the population dynamics. The particularity of QGA stems from the quantum representation they adopt which allows representing the superposition of all potential solutions for a given problem. It also stems from the quantum operators it uses to evolve the entire population through generations.

4. The Proposed Approach

In order to show how QC concepts have been tailored to the problem at hand, a formulation of the problem in term of quantum representation was derived and a quantum genetic dynamics borrowing quantum operations was defined (Figure 1). Then, we describe how these defined concepts have been integrated in a genetic algorithm.

4.1 Quantum Representation of Alignment

To successfully apply quantum principles on multiple sequence alignment, we have needed to map potential solutions into a quantum representation that could be easily manipulated by quantum operators. The multiple sequence alignment $S' = \{s'_1, s'_2, ..., s'_n\}$ is viewed as a binary matrix BM (Figure 2) where:

1. Each row represent a sequence s_i in the alignment
2. The value 1 denotes the presence of a basis that is a letter in alphabet Λ and the value 0 denotes the presence of a gap
3. The number of 1 in each line i is equal to the number of basis in the corresponding sequence

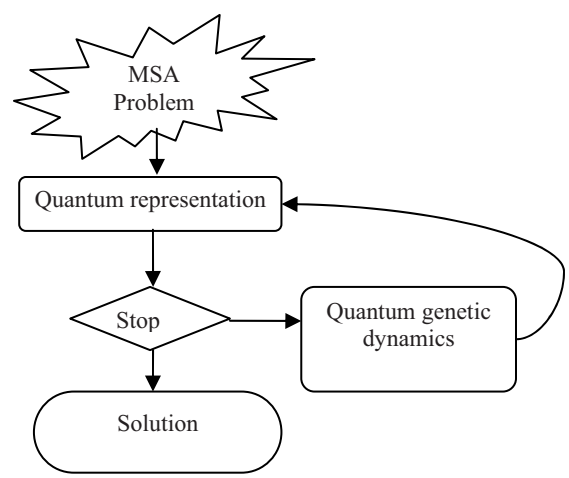

Figure 1. The core of the quantum genetic for the problem of multiple sequence alignment.

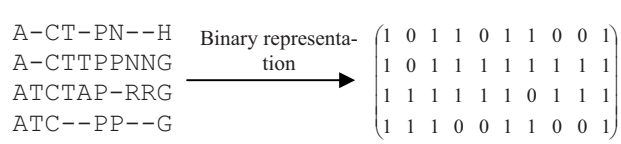

Figure 2. Binary representation of multiple alignment.

In term of quantum computing each sequence is represented as a quantum register as shown in figure 3. Each column $\begin{pmatrix} a_i \\ b_i \end{pmatrix}$ represents only one qubit and corresponds to an element of the alphabet Λ'. The amplitudes a_i and b_i are real values satisfying $|a_i|^2 + |b_i|^2 = 1$. For each qubit, a binary value is calculated according to the probabilities $|a_i|^2$, $|b_i|^2$ and the number of letters in each sequence. $|a_i|^2$ and $|b_i|^2$ are interpreted as the probabilities to have respectively an element of Λ or a gap. By consequence, all potential alignments can be represented by a quantum matrix QM (Figure 4) which contains a superposition of all possible configurations. This quantum matrix can be seen like a probabilistic representation of all possible alignments. This representation is effective when a genetic approach is adapted. It plays the role of the chromosome and it contributes to the reduction of the population size. Only one chromosome is needed to represent the entire population, i.e. all the solutions exist within each chromosome and what change are the probabilities to have one of them as a result of a measurement.

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_m \\ b_1 & b_2 & & b_m \end{pmatrix}$$

Figure 3. Quantum register representing a sequence.

$$\begin{bmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ b_{11} & b_{12} & & b_{1m} \end{pmatrix} \\ \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2m} \\ b_{21} & b_{22} & & b_{2m} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} a_{n1} & a_{n2} & \cdots & a_{nm} \\ b_{n1} & b_{n2} & & b_{nm} \end{pmatrix} \end{bmatrix}$$

Figure 4. Quantum matrix representing a multiple sequence alignment.

4.2 Basic quantum operations

The quantum operations which are the basis of the quantum genetic dynamics are as follows:

The operation of measure: This operation transforms by projection the quantum matrix into a binary matrix BM. this operation allows to extract from the quantum matrix one solution among all those present in superposition without destroying all other configurations as it is done in pure quantum systems. That has the advantage of preserving the superposition for the following iterations knowing that we operate on traditional machines. The value of a qubit is calculated according to its probabilities $|a_i|^2, |b_i|^2$ and the number of basis in each sequence. The obtained binary matrix is then translated into an alphabetical matrix by respecting the order of appearance of basis in the sequences. For example the figure 5 shows the extraction of the binary solution by using the operation of measure. Figure 6 shows the translation of the binary matrix into alphabetical matrix which describes for example a possible alignment of sequences AGA, AA, AGCA, and AA.

Quantum interference: This operation increases the amplitude of the best solution and decreases amplitudes of bad solutions. In another way, this operation aims to increase the probability for a good alignment to be extracted as a result of the measurement operation. It mainly consists in moving the state of each qubit in the direction of the value of the bit corresponding in the current best solution. The operation of interference is useful to intensify research around the best solution. This operation can be accomplished by using a unit transformation which achieves a rotation whose angle is a function of the amplitudes a_i, b_i and of the value of the corresponding bit in the solution reference (Figure 7). The values of the rotation angle δθ is chosen so that to avoid premature convergence. It is set experimentally and its direction is determined as a function of the values of a_i, b_i and the corresponding element's value in the binary matrix as shown in Table 1.

$$\begin{bmatrix} \begin{pmatrix} 0.70 & 0.44 & -0.90 & 0.77 \\ 0.70 & 0.99 & 0.44 & 0.63 \end{pmatrix} \\ \begin{pmatrix} 0.77 & 0.77 & 0.70 & 0.90 \\ 0.60 & 0.63 & 0.70 & 0.44 \end{pmatrix} \\ \begin{pmatrix} 0.63 & 0.44 & 0.99 & 1.00 \\ 0.77 & 0.90 & -0.14 & 0.00 \end{pmatrix} \\ \begin{pmatrix} 0.70 & 0.63 & 0.77 & 0.99 \\ 0.70 & 0.77 & 0.63 & 0.14 \end{pmatrix} \end{bmatrix} \xrightarrow{\text{Measure}} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Figure 5. Measure of quantum matrix

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{TRADUCTION} \begin{pmatrix} A & - & G & A \\ - & A & - & A \\ A & G & C & A \\ - & - & A & A \end{pmatrix}$$

Figure 6. Translation of binary matrix into alphabetic matrix

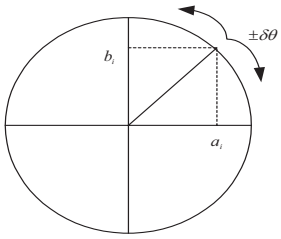

Figure 7. Quantum interference

Table 1. Lookup table of the rotation angle

a	*b*	Reference bit value	Angle
> 0	> 0	1	+δθ
> 0	> 0	0	-δθ
> 0	< 0	1	-δθ
> 0	< 0	0	+δθ
< 0	> 0	1	-δθ
< 0	> 0	0	+δθ
< 0	< 0	1	+δθ
< 0	< 0	0	-δθ

Mutation operators: These operators allow exploring new solutions and thus enhance the diversification capacities of the search process. They allow moving from the current solution to one of its neighbours. They consist mainly to alter some qubits in chromosomes. We have readapted and redefined some mutation operators mentioned in [6] for the quantum matrix alignment. In our algorithm we have used seven kinds of mutations which can be classified in three great classes:

- *The single qubit mutation*: in this mutation, we alter some qubits taken randomly (figure 8.a).
- *The register of qubits Mutation*: a random set of consecutive qubits are moved (figure 8.b).
- *The bloc of qubits mutation*: a random bloc of qubits is moved (figure 8.c).

Crossover operators: Crossovers are important for promoting the exchange of high quality blocks within the population. They exchange subparts of two quantum chromosomes. Taking inspiration from SAGA [6], we have defined two kinds of quantum crossovers (one-point and uniform) for the case of two quantum chromosomes. For example the figure 9 shows a quantum uniform crossover.

4.3. Outline of the proposed framework

Now, we describe how the representation scheme including quantum representation and quantum operators has been embedded within a genetic algorithm and resulted in a hybrid stochastic algorithm performing multiple sequence alignment. In the first stage, a population of 4 quantum chromosomes is created. Each chromosome contains a quantum alignment matrix. Then we compute

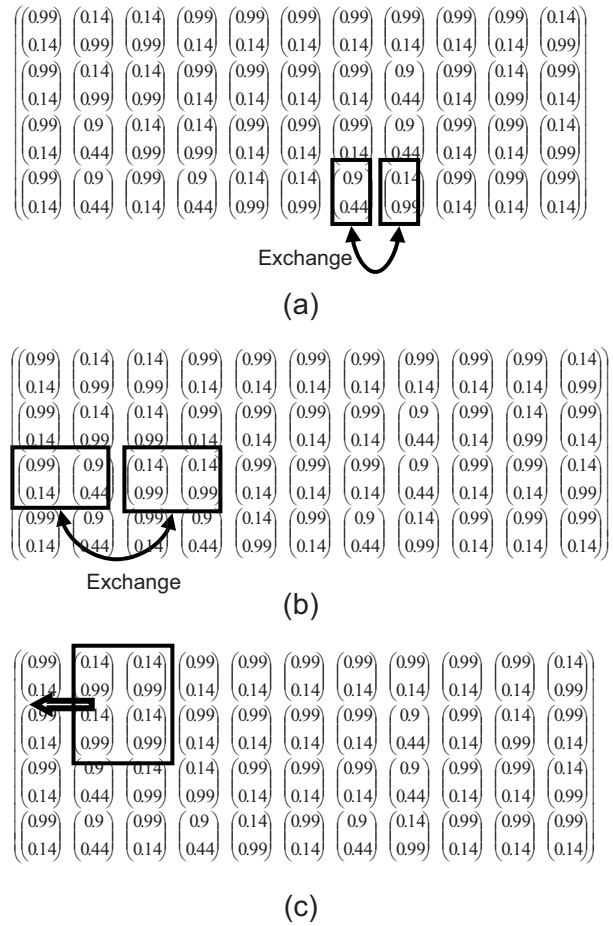

Figure 8. Mutation operators: (a) Single qubit mutation. (b) Register of qubits mutation. (c) Bloc of qubits mutation

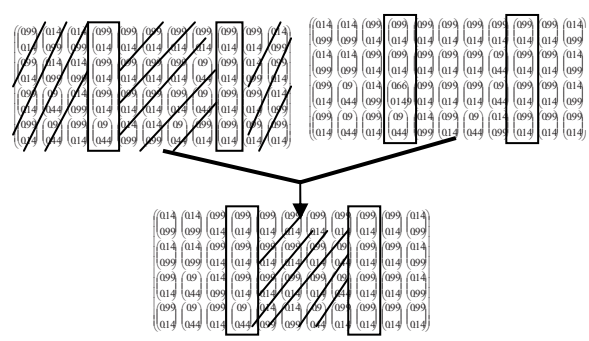

Figure 9. Quantum uniform crossover

an initial alignment by using the progressive method of Feng and Dollitlle [4]. First we compute the global pairwise alignment of all possible pairs of sequences, then the alignments scores are transformed into distances to build a guide tree. Finally we align each node of the guide tree according to the order given by the tree. This solution is translated into binary matrix. In each step of our program we apply firstly the interference operation. Secondly, the crossover operation is applied. Therefore, we obtain from the 4 initial chromosomes 12 new ones. The population becomes composed of 16 chromosomes. The crossover is followed by the quantum mutation; only one type of mutations is selected randomly. The fourth operation is the measurement operation which gives a binary matrix BM. This last is translated into alphabetic matrix according to the set of sequences. Lastly, we perform a selection of 4 chromosomes among the 16 existing in the current generation. We select the 3 chromosomes from which obtain the 3 best alignments and we select also arbitrarily one chromosome among the others in order to create a new population for the next generation. Every solution is evaluated by using the COFFEE function [12]. The global best solution is then updated if a better one is found and the whole process is repeated until having satisfaction of a stopping criterion. In more details, the proposed quantum genetic algorithm for multiple sequence alignment can be described as:

Input: A set of sequences SEQ

(1) Generate population of 4 chromosomes QPOP.
(2) Generate an initial alignment Aln_0. Let BM be the corresponding binary matrix.
(3) Set $Aln_{best} = Aln_0$ and $C_{best} = C(Aln_0)$.

Repeat

(4) Apply an interference operation according to the best solution.
(5) Apply a crossover operation.
(6) Apply a mutation operation.
(7) Apply a measurement operation on each chromosome to derive a new binary matrix BM_i.
(8) For each BM_i, evaluate the corresponding alignment Aln_i.
(9) If $C(Aln_{best}) < C(Aln_i)$ then $Aln_{best} = Aln_i$ and $C_{best} = C(Aln_i)$.

Until a termination-criterion is reached.

Output: Aln_{best} and $C(Aln_{best})$

4.5. Objective function

The objective function is used to measure the overall alignment quality of each chromosome. In this paper we have used COFFEE function [12] as fitness criterion. The basic idea of COFFEE is firstly to generate a library of all pairwise alignment of the sequences in the alignment. Secondly, it computes the level of identity between the current multiple alignment and the pairwise library. The COFFEE score of each alignment is calculated by the following formula:

$$Score_COFFEE = \frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} W_{ij} * Score(A_{ij})}{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} W_{ij} * Len} \quad (5)$$

Where N is the number of sequences to be aligned, A_{ij} is the pairwise projection of sequences S_i and S_j obtained from the multiple alignment, $Score(A_{ij})$ is the level of identity between A_{ij} and the corresponding pairwise alignment in the library. W_{ij} is the weight associated with two aligned sequences S_i and S_j and Len is the length of the multiple alignment [12]. When COFFEE is used as objective function, an optimal MSA solution is one which achieves the maximum COFFEE score.

5. Evaluation And Discussion

Our approach is implemented in MATLAB 7 tool and tested on home PC. To assess the efficiency and accuracy of our approach several experiments were designed. We have used the BALiBASE benchmark alignment database version 2 which contains manually refined multiple sequence alignments and biologically are valid [13]. BALiBASE is organized into five references sets: Reference 1 includes equidistant sequences of similar length with small insertions and extensions, it includes three subgroups V1, V2, V2; Reference 2 consists of closely-related sequences families with up to three distant "orphan" sequences; Refe-rence 3 contains equidistant divergent families; Reference 4 contains sequences with large N/C terminal extensions; and at last, Reference 5 contains sequences with large internal insertions [18]. The Core blocks in each alignment define the regions that can be unfailingly aligned excluding ambiguous or non-superimposable three-dimensional regions. It represents 58% of the residues in the alignments [18]. We analyzed the results of our approach by comparing the alignments produced by our algorithm with those obtained through other leading alignment techniques including CLUSTAL, DIALIGN and TCOFFEE. The results of these programs are collected by running their programs on our system. To estimate the biological quality of each alignment, we have used tow measures SPS and CS given by the program *Bali Score* and the core bloc annotation file which are available on the BALiBASE site. The Sum of Pairs Score SPS gives the percent of pairs correctly aligned while the Columns

Score CS gives the percent of columns correctly aligned. Moreover, we have performed another test in order to measure the accuracy of our approach in terms of secondary structure information. We have used 25 tests taken from the base of RNA benchmarks BRALIBASE [14]. It is divided into five references: Intron, rRNA, SRP, tRNA and U5. Two measures are used to measure the program quality. The first is the famous SPS which provides a measure of the sensitivity of the prediction [14]. The second measure is the structural conserved index SCI which gives the precision of the conserved secondary structure information contained within the alignment. The SCI is close to zero if there is no common RNA structure in the alignment. A set of perfectly conserved structures has an SCI≈1. An SCI>1 shows that there is a conserved RNA secondary structure which is, in addition, supported by compensatory and/or consistent mutations [19]. We have compared our results on BRALIBASE tests with the results of the programs Clustal, DIALIGN, MAFFT and PROALIGN. The results of these programs are taken from www.binf.ku.dk/users/pgardner/bralibase/.

Tests using BALiBASE benchmark data set display clearly the potent of using QGA to make global multiple sequence alignment (table 2). Our approach creates alignments with average accuracy comparable with or superior to the current methods. In the reference 1, our program outperforms all programs. In the rest of references, the average SPS is comparable to the other programs. However, in reference 4 and 5 the average CS is slightly below than that of DIALIGN and TCOFFEE.

By another hand, the table 3 summarizes the performance of different alignment methods including our approach, CLUSTAL, DIALIGN and PROALIGN upon structural RNA. The results state clearly that our approach can be useful in the structural prediction. In the most cases, it ranks high on SCI results.

6. Conclusion

The objective of this study is to validate the efficacy of quantum genetic algorithm and assess it as compared to other commonly used techniques for multiple sequence alignment. Compared to other techniques, the obtained results are encouraged. In the most cases, our program improves both the mathematical and the biological quality. In comparison to classical genetic algorithms, the proposed quantum genetic algorithm reduces efficiently the population size and the number of iterations to have the optimal solution. Thanks to superposition, interference, crossover and mutation operators, better balance between intensification and diversification of the search is achieved. However, there are several issues to improve our program. Firstly, the choice of the objective function is very important to have alignments of best quality, so, as future work, we will study the effects of using different objective functions. Secondly, to raise the speed of our program it's better to use parallels machines because it was verified effectively that QGA can work better on parallels machines. Thirdly, it's better to perform much refinement in poorly aligned blocs to reduce the runtime.

Table 2. Comparison of BALiBASE performance for our approach (OUR), DIALIGN (D), CLUSTAL (C) and TCOFFEE (T).

	SCORE	Ref1	Ref2	Ref3	Ref4	Ref5
OUR	SPS	0,890	0.917	0. 776	0.876	0.929
	SCI	0.832	0.585	0.468	0.684	0.818
D	SPS	0.811	0.893	0.684	0.897	0.940
	SCI	0.709	0.359	0.344	0.762	0.840
C	SPS	0.861	0.932	0.753	0.834	0.859
	SCI	0.773	0.568	0.460	0.522	0.638
T	SPS	0.866	0.934	0.785	0.918	0.958
	SCI	0.774	0.561	0.487	0.730	0.903

Table 3. Comparison of BRALIBASE performance for our approach (OUR), CLUSTAL (C), DIALIGN (D), MAFFT (M) and PROALIGN (P).

	SCORE	INTRON	rRNA	SRP	tRNA	U5
OUR	SPS	0,588	0,982	0,509	0,772	0,722
	SCI	0,660	0,904	0,738	1,052	0,674
C	SPS	0,611	0.990	0.292	0.858	0.793
	SCI	0.640	0.900	0.694	0.904	0.648
D	SPS	0.466	0.964	0.261	0.799	0.863
	SCI	0.602	0.832	0.668	0.752	0.616
M	SPS	0.543	0.947	0.231	0.788	0.830
	SCI	0.418	0.718	0.440	0.682	0.484
P	SPS	0.536	0.982	0.362	0.915	0.889
	SCI	0.596	0.900	0.732	1.022	0.688

Acknowledgment

This work was supported by CMEP-programme TASSILI under project 05 MDU 642.

References

[1] L. Wang and T. Jiang. On the complexity of multiple sequence alignment. J. Comput. Biol., volume 1, pages 337-348, 1994.

[2] C. Notredame. Recent progresses in MSA a survey. pharmacogenomic, volume 3, pages 1–14, 2002.

[3] S.B. Needlman and C.D. Wunsch. A general method applicable to the search for similarities in the amino-acid sequence of two proteins. Journal of Molecular Biology, volume 48, pages 443-453, 1970.

[4] D. Feng and R. Doolittle. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol., volume 25, pages 351-360, 1987.

[5] J. Kim , S. Pramanik and M.J. Chung. Multiple sequence alignment using simulated annealing. Computer applications in bioscience, volume 10, pages 419-426, 1994.

[6] C. Notredame and D.G. Higgins. SAGA: sequence alignment by genetic algorithm, Nucleic Acids Research, volume 24(8): 1515–1524, 1996.

[7] T. Riaz, Y. Wang and K.B. Li. Multiple sequence alignment using Tabu Search. Proc. 2nd Asia-Pacific Bioinformatics Conference (APBC), pages 223-232, 2004.

[8] K.-H. Han and J.-H. Kim. Genetic quantum algorithm and its application to combinatorial optimization problem. Proc. 2000 Congr. Evolutionary Computation, volume 2, La Jolla, CA, pages 1354–1360, 2000.

[9] K.-H. Han and J.-H. Kim. Quantum-inspired Evolutionary Algorithms with a New Termination Criterion, Hε Gate, and Two Phase Scheme. IEEE Transactions on Evolutionary Computation, IEEE Press, volume 8(2):156-169, April 2004.

[10] S. Meshoul, A. Layeb and M. Batouche. A Quantum Evolutionary Algorithm for effective Multiple Sequence Alignment. In the proc of 12th Portuguese conference on artificial intelligence (EPIA 2005 CMB workshop). Lecture notes in artificial intelligence (LNAI), volume 3808, pages 260-271, 2005.

[11] J. D. Thompson, F. Plewniak, and O. Poch. BALiBASE: A benchmark alignment database for the evaluation of multiple alignment programs, Bioinformatics, volume 15, pages 87-88, 1999.

[12] C. Notredame, L. Holm and D.G. Higgins. COFFEE: An objective functions for multiple sequence alignments. Bioinformatics, volume 14, No 5, pages 407-22, 1998.

[13] J.D. Thompson D. Higgins and T. Gibson. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting positionspecific gap penalties and weight matrix choice. Nucleic Acids Res. volume 22, pages 4673-4690, 1994.

[14] P. Gardner, A. Wilm and S. Washietl. A benchmark of multiple sequence alignment programs upon structural RNAs. Nucleic Acids Research, volume 33(8): 2433–2439, 2005.

[15] P.W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. Proc. 35th Annu. Symp. Foundations Computer Science, Sante Fe, NM, pages 124–134, 1994.

[16] L.K. Grover. A fast quantum mechanical algorithm for database search. Proc. 28th ACM Symp. Theory Computing, pages 212–219, 1996.

[17] C.P. Williams and S.H. Clearwater. Explorations in quantum computing. Springer Verlag, Berlin, Germany 1998.

[18] J. D. Thompson, F. Plewniak, and O. Poch. A comprehensive comparison of multiple sequence alignment programs, Nucleic Acids Research, volume 27(13):2682-2690, 1999.

[19] S. Washietl, I. Hofacker and P. Stadler. Fast and reliable prediction of noncoding RNAs. Proc. Natl Acad. Sci. volume 102, USA, pages 2454–2459, 2005.