



RESEARCH ARTICLE

INHERIT BIOLOGICAL CHARACTERISTICS TO OPTIMIZE GENETIC MULTIPLE SEQUENCE ALIGNMENT USING HYBRID META-HEURISTIC MACHINE LEARNING APPROACH

¹*Heena Arora and ²Avani Chopra

¹Researcher, Computer Science and Engineering, D.A.V.I.T., Jalandhar, India

²Asst. Professor, Information Technology, D.A.V.I.T., Jalandhar, India

Received 26th May, 2018; Accepted 20th June, 2018; Published 30th July, 2018

ABSTRACT

Traditionally Optimization of Multiple Genetic Sequence Alignment is based on statistics algorithms, which leads to loss of genetic characteristics and generate less feasible results. Sequence Alignment of genetic sequences need qualitative characteristics configuration at each level of MSA iteration, especially where characteristics inheritance based on specially genomes properties. Multiple Sequence Alignment optimization is one of the prominent areas in Bioinformatics where optimization problems occur due to mutation, uncertainty of large genetic data sets. Data Science and Mathematical Optimization Algorithm provides the combine solution for optimization problems related to genetic alignment. This research paper proposed a novel Hybrid Meta-Heuristic Optimization Genetic Alignment Algorithm (HMHOGA) to solve genetic alignment optimization problem at each phase of MSA iteration and also help to recognize occurrence of meta-heuristic characteristics. Proposed algorithm provides optimize alignment along with various evaluation factors like pair_sum(q), total_col_score(tc), standard_deviation_of_pair_matrix(z), reliability(r) and efficiency(e) based on CPU execution time. These evaluation factors act like as meta-heuristic knowledge base to optimize future alignments.

Keywords: Bioinformatics, ClustalW, ClustalOmega, MUSCLE, Multiple Sequence Alignment, Optimization, Hybrid Meta-heuristic, Machine Learning.

Copyright © 2018, Heena Arora and Avani Chopra. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Heena Arora and Avani Chopra, 2018. "Inherit biological characteristics to optimize genetic multiple sequence alignment using hybrid meta-heuristic machine learning approach" *International Journal of Current Research in Life Sciences*, 7, (07), 2513-2520

INTRODUCTION

Intricate evolution of genetic species existence comes due to trillion years of mutation. This genetic mutation has been continuously generated large amount of labyrinthine data sets, which need to be optimized in field of bioinformatics. Genetic information and its mutation generate knowledge gap, which need to be fill with intelligent scientific approaches. Major challenge of native era is to evolve optimal machine learning and data science algorithms to solve these problems. In living beings Chromosome to Protein generation is complex mutant structure at each level, these types of mutations generate various infected cells, which leads towards critical diseases, so to track these types of mutation, we need to generate algorithms with the help of data science, which helps to track diseases mutation and to find its remedies for living beings. Alvaro *et al.*, (2015), proposed the algorithm which uses the Evolutionary Computation and Multi objective Optimization for solving this bioinformatics problem. They represented a multi objective version of a memetic metaheuristic called Hybrid Multi objective Metaheuristics.

The results obtained by their method are compared with well-known methods. At the end they concluded that their new approach presents remarkable accuracy when it deals with the sequences which have a low sequence similarity. The study presented by Alvaro Rubio-Largo proposed a characteristic-based framework for improving the accuracy of MSA by characteristics of the input set of unaligned sequences, which need to extend search space for efficient optimize results with the help of meta-heuristic approach. Alvaro *et al.*, (2016), proposed a characteristic based framework for multiple sequence aligners. In this they used input set of unaligned sequences than extract its characteristics aligner According to the best parameter configuration they run the aligner. At the end they show that their proposed work is better than default muscle by compare this framework with using different benchmarks. The results show that the alignment of higher accuracy is achieved by using proposed framework. Alvaro Rubio-Largo also proposed a memetic algorithm for sequence alignment, which works better in local search space and need requirement to extend the search space. Anderson *et al.*, (2016), developed a parallel version of the MSA-GA tool which is based on genetic Algorithms approach using multithread programming, their results shown that the modification version is 11% faster than the standard MSA-GA, using WSP function. Their results shown that their proposed

*Corresponding author: Heena Arora,
Researcher, Computer Science and Engineering, D.A.V.I.T., Jalandhar, India.

work good results produced good results and improved execution time. Dengfeng *et al.*, (2015), found consensus sequence. For compare each sequence in the consensus sequence they used simulated annealing and star alignment. In this they present a multiple sequence alignment method based on the determined consensus sequence. Their Experimental results showed high accuracy of the algorithm. Huazheng Z. *et al.*, (2016), proposed algorithm called MOMSA in which they used a new population initialization method and a novel mutation operator. They compared the performance of MOMSA with several alignment techniques including IMSA, MSAProbs, Probalign, MAFFT, Clustal omega, T-Coffee, Kalign2, MUSCLE, CLUSTALW. For testing they had used benchmark datasets BALiBASE 2.0 and BALiBASE 3.0. Their results shown that their proposed algorithm called MOMSA obtained better alignments than VDGA, GAPAM on the most of test cases VDGA, GAPAM on the most of test cases, and in in terms of TC scores it gave better alignments than IMSA. Alam Manaar *et al.* (2016) have found a parallel search technique that can quickly and efficiently solve several optimization problems by using very less tuneable parameters for solving fast and efficient as compared to similar stochastic algorithms. This approach uses rotating and complement operator to reach distinct nodes, instead of traversing repeated nodes in the search space and flip operator use to capture variations within the search space as well. In bioinformatics, Multiple sequence alignment is complex bioinformatics problem that involves the alignment of three or more genetic sequences. Multiple Sequence Alignment is generally used to match biological facts about genomes and proteins which is typical used to check mutations. MSA also used to create a phylogenetic tree. Above mentioned reports and comparison with traditional approaches shows that our proposed hybrid meta-heuristic algorithm is generate better results and feasible approach to overcome these problems.

MATERIALS AND METHODS

Involvement of different characteristics sequences in Multiple Sequence Alignment generates Knowledge Gap based on various factors and properties. To fill this knowledge gap, biological characteristics matching of genetic sequences is best approach. Characteristics based MUSCLE algorithm is fit for this approach due to flexible configuration based on genetic characteristics. To inherit best configuration, first of all need to compare input sequences with references data set like BaliBASE v3.0, because its tested and verify manually by various scientific communities. At initial level BaliBASE v3.0 work as base and after execution of proposed algorithm native heuristic knowledge base inherit with reference dataset. To find optimal configuration our proposed system uses power of hybrid meta-heuristic machine learning, which learn from experience and increase knowledge base with every execution.

Theoretical Framework (Hybrid Meta-Heuristic Approach): Multiple Sequence Alignment based on mutation characteristics generated optimization issues, which is a very complex NP-Hard Problem and needs to be solved by scientific approach. Traditionally these problems solved by statistical algorithms, but due to involvement of various characteristics in genetic sequences need to more specific in terms of biological characteristics of genetic sequences and proposed a suitable solution in form of native genetic optimize aligner based on Hybrid Meta-Heuristic Approach. The existing framework has been designed for the characteristics-

based sequence alignment, especially for amino acid chain sequences. This framework helps to track mutation of similar characteristics proteins. This approach also helpful to predict and track various diseases cells in different gene structure and it's causes, which helps to develop remedies for living beings. This framework is very suitable for local search space, especially in the case of translation from mRNA to protein sequences and also useful for nucleic acid alignments.

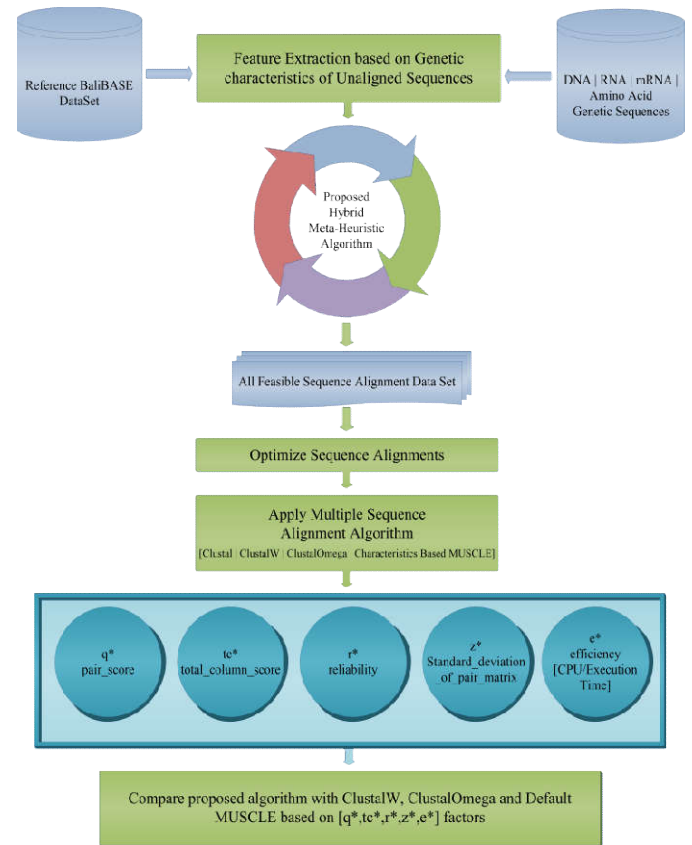


Figure 1. Block Diagram of Proposed Meta-Heuristic MUSCLE Algorithm

Proposed Heuristic & Meta-Heuristic Functions

a) q^* [Pair-Wise Alignment Score]

Pair-wise Alignment Score for quality measurement reference from Alvaro Rubio-Largo [3].

$$q^*(x) \leftarrow \sum_{k=0}^n \frac{qi}{qri} \quad (1)$$

b) tc^* [Total Column Score]

Total Column Score for quality measurement reference from Alvaro Rubio-Largo [3].

$$tc^*(x) \leftarrow \frac{\sum_{k=0}^n Ci}{n} \quad (2)$$

c) z^* [Standard Deviation of Similarity Matrix]

Stand Deviation of Similarity Matrix for quality measurement reference from Align Stat [14].

$$z^*(x) \leftarrow \text{standard_deviation}(\text{similarity}_{matrix}) \quad (3)$$

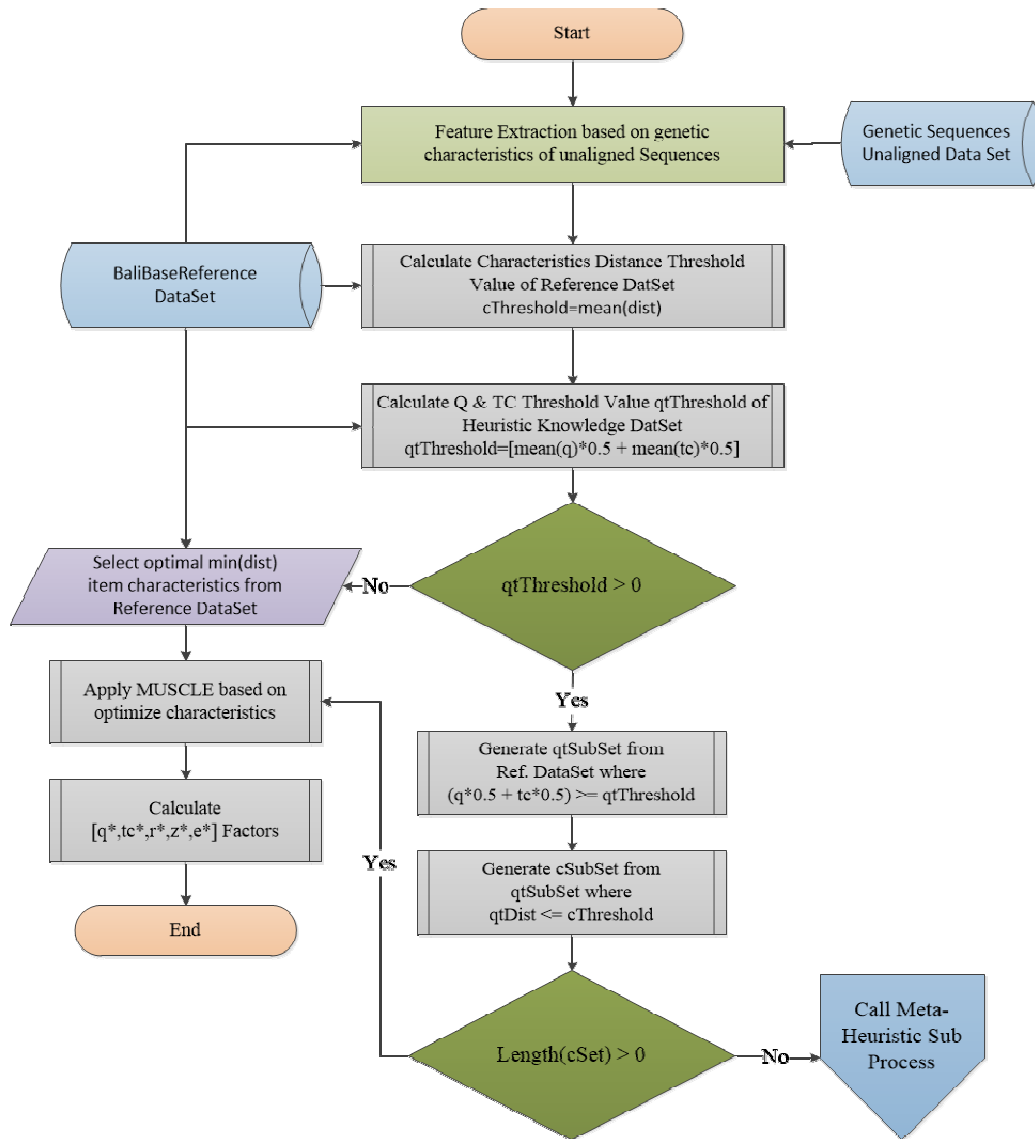


Figure 2. Purposed Hybrid Meta-Heuristic Algorithm (Heuristic Sub Process)

d) e* [Efficiency based on CPU Execution Time]

Efficiency value e* based on CPU execution Time of MSA algorithms [2].

e) r* [Reliability based on level of Similarity]

Reliability of MSA algorithm shows level of Similarity for quality measurement reference from Align Stat [14]. Here α is dissimilarity score.

$$r^*(x) \leftarrow 1 - \alpha \tag{4}$$

Proposed Hybrid Meta-Heuristic Optimization Algorithm

Algorithm:Proposed Hybrid Meta-Heuristic Optimization Algorithm to find optimal characteristics of MUSCLE for best optimal multiple sequence alignment.

Input: An Input Set of unaligned Sequences X and BaliBase v3.0 0 References DataSet B.

Output: Optimal MUSCLE Aligned Sequences X* and evaluation factors q*, tc*, e*, z* and r*.

Step 1: Start.

Step 2:Feature extraction of unaligned input sequences X using characteristics based fwkMUSCLE Algorithm 0.

Step 3: Call proposed Hybrid Meta-Heuristic (Heuristic Sub Process) to generate optimal characteristics C* hybridMetaHeuristicMUSCLE MSA along with evaluation factors q*, tc*, e*, z* and r*.

Step 4:Apply ClustalW, ClustalOmega and Default MUSCLE on input sequences and generate evaluation factors q*, tc*, e*, z* and r* for all applied MSA algorithms.

Step 5: Compare proposed algorithm with ClustalW, ClustalOmega and Default MUSCLE based on evaluation factors q*, tc*, e*, z* and r*.

Step 6: Stop.

Proposed Hybrid Meta-Heuristic Optimization Algorithm (Heuristic Sub Process)

Algorithm: Proposed Hybrid Meta-Heuristic Optimization Algorithm Heuristic Sub Process to find optimal characteristics of MUSCLE for best optimal multiple sequence alignment.

Input: An Input Set of unaligned Sequences X, unaligned sequences characteristics X(seqCharacteristics) and BaliBase v3.0 0 References Data Set with Meta-Heuristic Knowledge Base B*.

Output: Optimal MUSCLE Aligned Sequences X* and evaluation factors q*, tc*, e*, z* and r*.

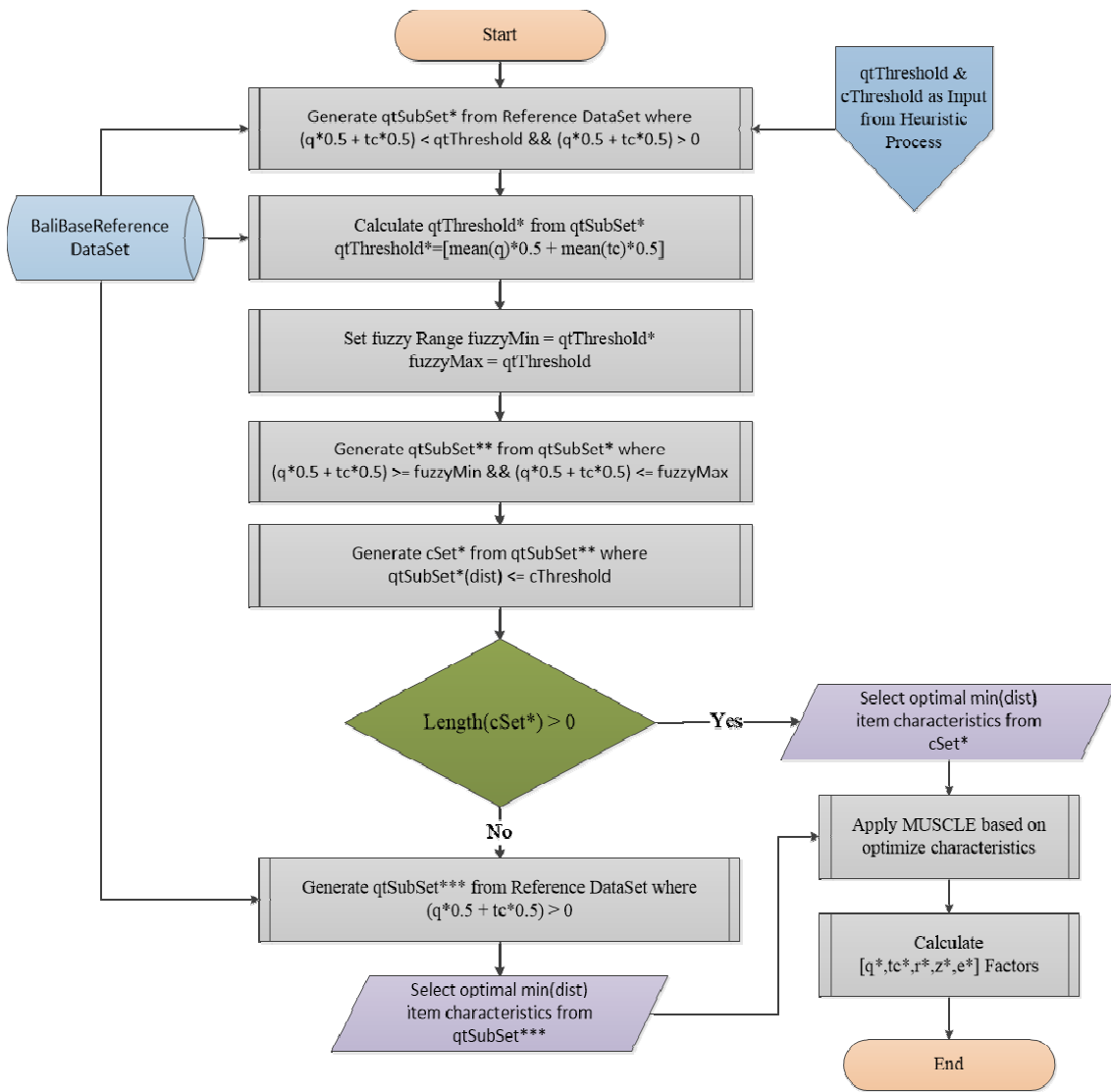


Figure 3: Purposed Hybrid Meta-Heuristic Algorithm (Meta-Heuristic Sub Process)

```

>1aab_
GKGDPKKPRGKMSSYAFFVQTSREEHKKKHPDASVNFSEFSKCCSERWKTMSAKEKGFEDMAKADKARYEREMKTYIPPKGE
>1j46_A
MQDRVKRPMAFIVWSRDQRRKMALENPRMRNSEISKQLGYQWKMLTEAEKWPFQEAQKLQAMHREKYPNYKYP RRKAKMLPK
>1k99_A
MKKLLKHPDFPKKPLTPYFRFFMEKRAKYAKLHPEMSNDLTKILSKKYKELPEKMKMYIQDFQREKQEFERNLARFREDHPDLIQNAKK
>21ef_A
MHIKKPLNAFMLYMKEMRANVVAESTLKE SAAINQILGRRWHALSREEQAKYYELARKERQLHMQLYPGWSARDNYGKKKKRREK
    
```

Figure 4. Unaligned Sequences of BaliBASE Benchmark BB11001

```

|>seq1
-----MQDRVKRPMAFIVWSRDQRRKMALENPRMR-NS-EISKQLGYQWKMLTEAEKWPFQEAQKLQ
AMHREKYPNYKYP RRKAKMLPK---
>seq2
MKKLLKHPDFPKKPLTPYFRFFMEKRAKYAKLHPEMS-NL-DLTKILSKKYKELPEKMKMYIQDFQREK
QEFERNLA--RFREDHPDLIQNAKK-
>seq3
---GKGDPKKPRGKMSSYAFFVQTSREEHKKKHPDASVNFSEFSKCCSERWKTMSAKEKGFEDMAKADK
ARYEREMK--TYIPPKGE-----
>seq4
-----MHIKKPLNAFMLYMKEMRANVVAESTLKE-SA-AINQILGRRWHALSREEQAKYYELARKER
QLHMQLYPGWSARDNYGKKKKRREK
    
```

Figure 5. Proposed Hybrid Meta-Heuristic MUSCLE MSA FASTA Output

Table 1. Evaluation Factors For ClustalW and Proposed Hybrid Meta-Heuristic MUSCLE

#	Evaluation Factor	Optimal Score(Probability)	Optimal Score(%)
1	$q^*(x) \leftarrow \sum_{k=0}^n \frac{qi}{qri}$	0.8744	87.44%
2	$tc^*(x) \leftarrow \frac{\sum_{k=0}^n Ci}{n}$	0.6875	68.75%
3	$z^*(x) = sd(smatrix)$	0.0895	8.95%
4	$e^*(cpu_time)$ [seconds]	4.8225	-
5	$r^*(x) \leftarrow 1 - \alpha$	0.8646	86.46%

Table 2. Evaluation Factors For ClustalOmega & Proposed Hybrid Meta-Heuristic MUSCLE

#	Evaluation Factor	Optimal Score(Probability)	Optimal Score(%)
1	$q^*(x) \leftarrow \sum_{k=0}^n \frac{qi}{qri}$	0.8312	83.12%
2	$tc^*(x) \leftarrow \frac{\sum_{k=0}^n Ci}{n}$	0.5625	56.25%
3	$z^*(x) = sd(smatrix)$	0.0879	8.79%
4	$e^*(cpu_time)$ [seconds]	6.6604	-
5	$r^*(x) \leftarrow 1 - \alpha$	0.8845	88.45%



Figure 6. Proposed Hybrid Meta-Heuristic MUSCLE MSA Pretty Print Output

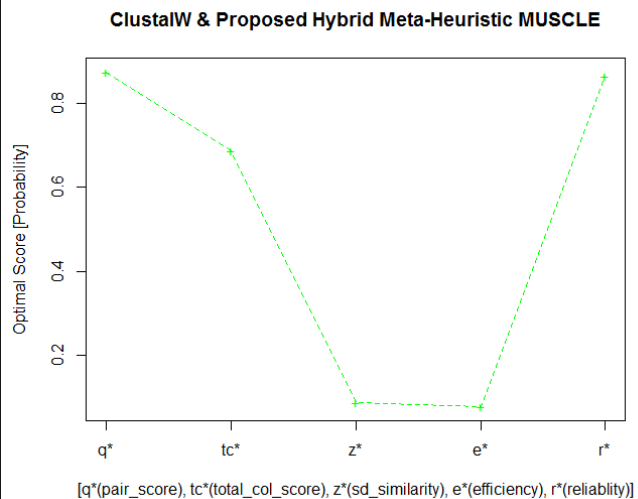
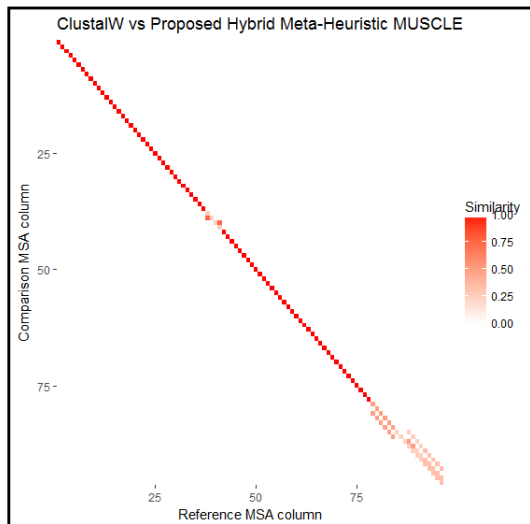


Figure 7. ClustalW & Proposed Hybrid Meta-Heuristic Figure 8. ClustalW & Proposed Hybrid Meta-Heuristic

```

>seq1
-----MQDRVVRKPMNAFIVWSRDQRRKMALENPRMRN--SEISKQLGYQWKMLTEAEKWPFFQEAQKLQ
AMHREKYPNYKYRP---RRKAKMLPK
>seq2
-----MHIKKPLNAFMLYMKEMRANVVAESTLKES--AAINQILGRRWHALSREEQAKYELARKER
QLHMQLYPGWSARDNYGKKKKRREK
>seq3
MKKLLKHPDFPKKPLTPYFRFFMEKRKYAKLHPMSN--LDLTKILSKYKELPEKKMKYIQDFQREK
QEFERNLARFREDH---PDLIQNAKK
>seq4
---GKGDPKKPRGKMSSYAFFVQTSREEHKKHPDASVNFSEFSKKSERWKTMSAKEGKFEDMAKADK
ARYEREMKTYIPPK---GE-----
    
```

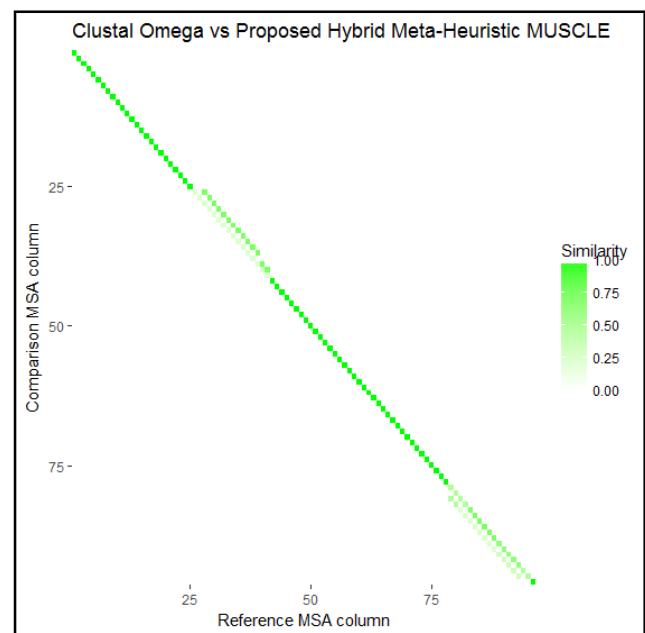
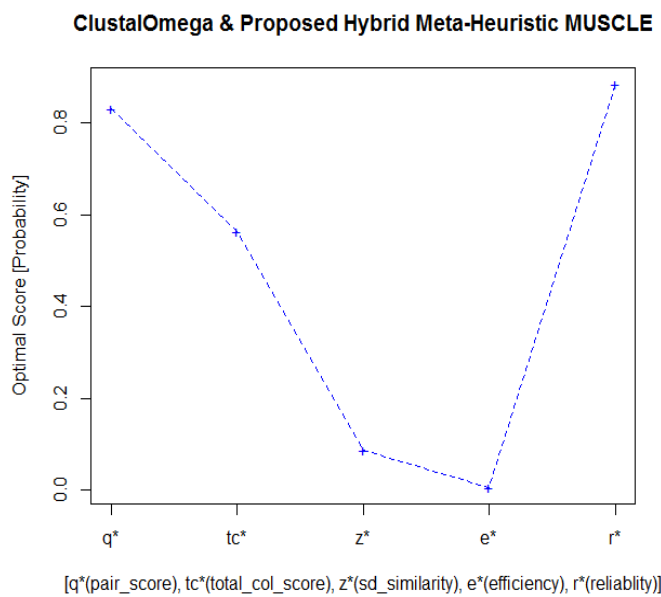
Figure 9. ClustalW MSA FASTA Output

Table 3. Evaluation Factors For Default MUSCLE & Proposed Hybrid Meta-Heuristic MUSCLE

#	Evaluation Factor	Optimal Score(Probability)	Optimal Score(%)
1	$q^*(x) \leftarrow \sum_{k=0}^n \frac{qi}{qri}$	0.9752	97.52%
2	$tc^*(x) \leftarrow \frac{\sum_{k=0}^n Ci}{n}$	0.7500	75.00%
3	$z^*(x) = sd(smatrix)$	0.0930	9.30%
4	$e^*(cpu_time)$ [seconds]	2.7806	-
5	$r^*(x) \leftarrow 1 - \alpha$	0.9715	97.15%

Table 4. Comparison between proposed Hybrid Meta-Heuristic vs ClustalW vs ClustalOmega

#	MSA	Sum of Pair Score[q*]	Total Column Score[tc*]	Standard Deviation of Similarity Matrix[z*]	Execution Time (Efficiency) (seconds)[e*]	Reliability [r*]
1	Proposed Hybrid Meta-Heuristic	0.97520	0.7500	0.0930319	2.7806514	0.971
2	ClustalW	0.87448	0.6875	0.0895399	4.8225229	0.864
3	ClustalOmega	0.83127	0.5625	0.0879348	6.6604481	0.884

**Figure 10. ClustalOmega and Proposed Hybrid Meta-Heuristic** **Figure 11. ClustalOmega and Proposed Hybrid Meta-Heuristic**

Step 1: Start.

Step 2: Calculate distance between X(seqCharacteristics) from all B*(characteristics) and merge computed distance set with B*.

Step 3: Calculate characteristics distance threshold value of reference data set B*.

$$cThreshold = \frac{\sum_{k=0}^n B^*(dist)}{n}; n = 0,1,2, \dots, len(B^*)$$

Step 4: Calculate Q & TC threshold value qtThreshold of Heuristic knowledge data set B*.

$$qtThreshold = \left[\left(\frac{\sum_{k=0}^n B^*(q)}{n} \right) * 0.5 + \left(\frac{\sum_{k=0}^n B^*(tc)}{n} \right) * 0.5 \right]; n = 0,1,2, \dots, len(B^*)$$

Step 5: If qtThreshold > 0 than move to step 9 else continue with step 6.

Step 6: Generate qtSubSet from B* where [B*(q) * 0.5 + B*(tc) * 0.5] >= qtThreshold.

Step 7: Generate cSubSet from qtSubSet where qtSubSet(dist) <= cThreshold.

Step 8: If length(cSubSet) > 0 than move to step 10 else call Meta-Heuristic Sub Process.

Step 9: Select optimal min(distance) item characteristics C* from B* and move to step 11.

Step 10: Select optimal min(distance) item characteristics C* from cSubSet.

Step 11: Apply MUSCLE 0 based on optimized characteristics C*.

Step 12: Calculate evaluation factors q*, tc*, z*, e* and r*.

Step 13: Return evaluation factors for comparison with other MSA algorithms.

Step 14: Stop.

Proposed Hybrid Meta-Heuristic Optimization Algorithm (Meta-Heuristic Sub Process)

Algorithm: Proposed Hybrid Meta-Heuristic Optimization Algorithm Meta-Heuristic Sub Process to find optimal characteristics of MUSCLE for best optimal multiple sequence alignment.

Input: BaliBase v3.0 0 References Data Set with Meta-Heuristic Knowledge Base B*, qtThreshold and cThreshold values from Heuristic Sub Process.

Output: Optimal MUSCLE Aligned Sequences X* and evaluation factors q*, tc*, e*, z* and r*.

Step 1: Start.

Step 2: Generate qtSubSet* from B* where $[(B^*(q) * 0.5 + B^*(tc) * 0.5) < qt \text{ Threshold} \ \&\& \ (B^*(q) * 0.5 + B^*(tc) * 0.5) > 0]$.

Step 3: Calculate extended Q & TC threshold value qtThreshold* from qtSubSet*.

$$qtThreshold^* = \left[\left(\frac{\sum_{k=0}^n qtSubSet^*(q)}{n} \right) * 0.5 + \left(\frac{\sum_{k=0}^n qtSubSet^*(tc)}{n} \right) * 0.5 \right];$$

$$n = 0, 1, 2, \dots, len(qtSubSet^*)$$

Step 4: Set fuzzy range fuzzyMin = qtThreshold* and fuzzyMax = qtThreshold.

Step 5: Generate qtSubSet** from qtSubSet* where $[(qtSubSet^*(q) * 0.5 + qtSubSet^*(tc) * 0.5) \geq fuzzyMin \ \&\& \ (qtSubSet^*(q) * 0.5 + qtSubSet^*(tc) * 0.5) \leq fuzzyMax]$.

Step 6: Generate cSubSet* from qtSubSet** where qtSubSet**(dist) <= cThreshold.

Step 7: If length(cSubSet*) > 0 than move to step 10 else continue with step 8.

Step 8: Generate qtSubSet*** from B* where $[B^*(q) * 0.5 + B^*(tc) * 0.5] > 0$.

Step 9: Select optimal min(distance) item characteristics C* from qtSubSet*** and move to step 11.

Step 10: Select optimal min(distance) item characteristics C* from cSubSet*.

Step 11: Apply MUSCLE 0 based on optimized characteristics C*.

Step 12: Calculate evaluation factors q*, tc*, z*, e* and r*.

Step 13: Return evaluation factors for comparison with other MSA algorithms.

Step 14: Stop.

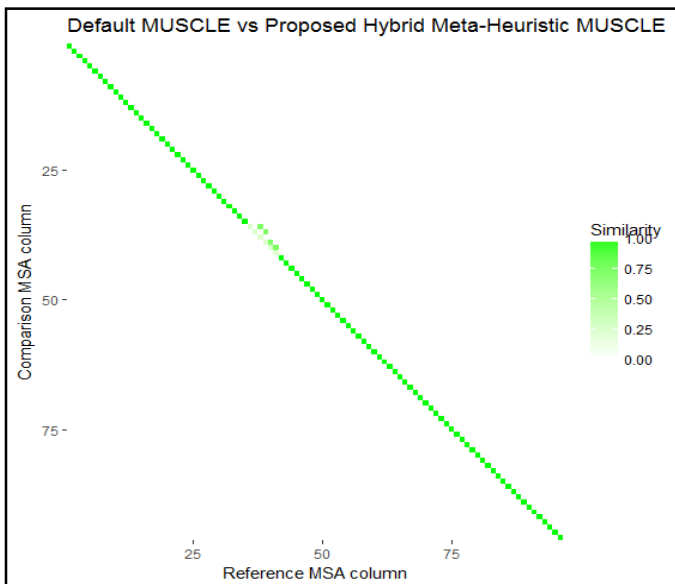


Figure 13. Default MUSCLE & Proposed Hybrid Meta-Heuristic

```

>seq1
-----MQDRVKRPMNAFIVWSRDQRKMALENPR--MRNSEISKQLGYQWKMLTEAEKWPFQEAQKIQ
AMHREKYPNYKYRPRRKAKMLPK---
>seq2
MKKLLKHPDFPKKPLTPYFRFFMEKRAKYAKLHPE--MSNLDLTKILSKKYKELPEKKMKMYIQDFQREK
QEFERNLA--RFREDHPDLIQNAKK-
>seq3
---GKGDPKKPRGKMSSYAFFVQTSREEHKKKHPDASVNFSEFSKCCSERWKTMSAKEKGFEDMAKADK
ARYEREMK--TYIPPKGE-----
>seq4
-----MHIKKPLNAFMLYMKEMRANVVAESTL--KESAAINQILGRRWHALSREEQAKYYELARKER
QLHMQLYPGWSARDNYGKKKKRKRK
    
```

Figure 14. Default MUSCLE MSA FASTA Output

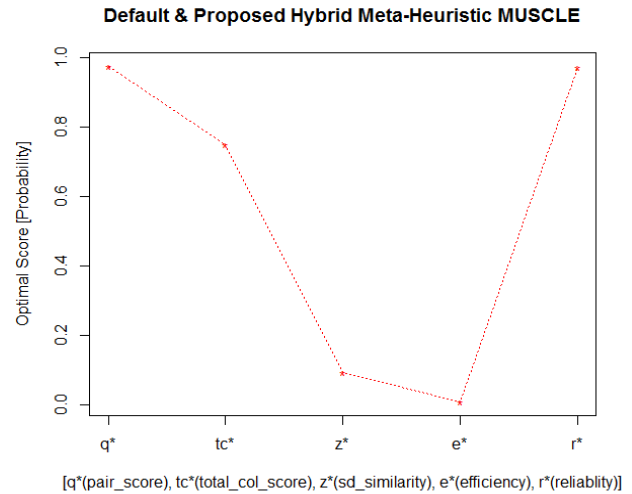


Figure 15. Default MUSCLE & Proposed Hybrid Meta-Heuristic

RESULTS AND DISCUSSION

Unaligned Sequences of Input Data Set (X): Unaligned Sequences of BaliBASE Benchmark BB11001 [9] input sequences used to perform multiple sequence alignment based on various algorithms.

Output Alignment Results (Y*) of Proposed Hybrid Meta-Heuristic MUSCLE: The result of proposed Hybrid Meta-Heuristic MUSCLE Algorithm.

Results(Y*) of ClustalW & Proposed Hybrid Meta-Heuristic MUSCLE: The result of cluatalW and proposed Hybrid Meta-Heruistic MUSCLE Algorithm.

Results(Y*) of ClustalOmega & Proposed Hybrid Meta-Heuristic MUSCLE: The result of cluatalOmega and proposed Hybrid Meta-Heruistic MUSCLE Algorithm.

Results(Y*) of Default MUSCLE & Proposed Hybrid Meta-Heuristic MUSCLE: The result of default MUSCLE and proposed Hybrid Meta-Heruistic MUSCLE Algorithm.

Comparison between Proposed Hybrid Meta-Heuristic vs ClustalW vs ClustalOmega

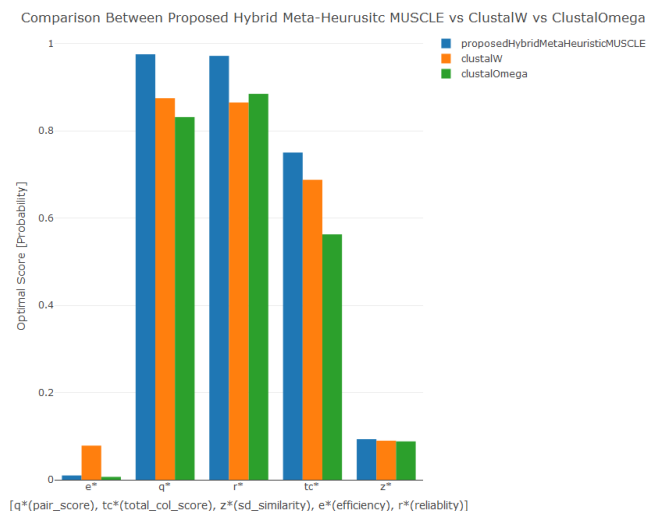


Figure 16. Comparison between Proposed Hybrid Meta-Heuristic MUSCLE vs ClustalW vs ClustalOmega

Conclusion

Genetic MSA algorithms are based on traditional approaches, which have lack of biological comparison using genetic characteristics, those are actual base of genetic mutation between different sequences. Characteristics based MUSCLE Algorithm is best to overcome this problem, because MUSCLE configuration provides various genetic benefits, but to choose best configuration need to create best optimization algorithm with intelligence and scientific approach. The proposed hybrid meta-heuristic optimization algorithm provide solution to overcome this problem and also increase knowledge base with the help machine learning. This proposed system generates best result for similar characteristics sequences and it's also work better for different characteristics sequences by comparison various judgement factors those are able to modify in future.

REFERENCES

- Alam Manaar, Chatterjee Soumyajit, Banka Haider, 2016. "A novel parallel search technique for optimization" in Proceedings of the 3rd International Conference on Recent Advances in Information Technology (RAIT), (Dhanbad, India), ISBN 978-1-4799-8579-1, vol. 1, pp. 259-263.
- Álvaro R. L., Leonardo V., Mauro C., and Miguel A. V.R. (2015) "A Hybrid Multi-Objective Memetic Metaheuristic for Multiple Sequence Alignment". IEEE Transaction on Evolutionary Computation, vol. 20, pp. 499 - 514.
- Álvaro R. L., Leonardo V., Mauro C., and Miguel A. V.R. 2016. "A Characteristic-Based Framework for Multiple Sequence Aligners". IEEE Transaction on Cybernetics, vol. 48, pp. 41-51.
- Anderson R. A. *et al.*, 2016. "Performance Improvement of Genetic Algorithm for Multiple Sequence Alignment". 17th International Conference on Parallel and Distributed Computing, Applications and Technologies, Guangzhou, China, pp. 70-72.
- Bacon D. J. and W. F. Anderson, 1986. "Multiple sequence alignment," *J. Mol. Biol.*, vol. 191, no. 2, pp. 153-161.
- Dengfeng Y., Minghu J. J., Xu. Y., Abudoukelimu A., Renkui H. 2015. "An Algorithm of Multiple Sequence Alignment Based on Consensus Sequence Searched by Simulated Annealing and Star Alignment". International Symposium on Bioelectronics and Bioinformatics (ISBB), Beijing, China, pp. 3-6.
- Edgar R. 2015. Muscle. [Refer Source Online]. Available: <http://www.drive5.com/muscle/>
- Edgar, R. C. 2004. "MUSCLE: Multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Res.*, vol. 32, no. 5, pp. 1792-1797.
- Huazheng Z., Zhongshi H., and Yuanyuan J. 2016. "A Novel Approach to Multiple Sequence Alignment Using Multi-Objective Evolutionary Algorithm Based on Decomposition". *IEEE Journal of Biomedical and Health Informatics*, vol.20, pp.717-727.
- Jordehi A. R. and J. Jasni, 2015. "Particle swarm optimisation for discrete optimisation problems: A review," *Artif. Intell. Rev.*, vol. 43, no. 2, pp. 243-258.
- Konstantina K., Costas P., and Dimitrios I. F. 2016. "Integration of Pathway Knowledge and Dynamic Bayesian Networks for the Prediction of Oral Cancer Recurrence". *IEEE Journal of Biomedical and Health Informatics*, vol.21, pp. 320-327.
- Liu Ruoqian, Agrawal Ankit, Liao Wei-keng, Choudhary Alok, Chen Zhengzhang, 2015. "Pruned search: A machine learning based meta-heuristic approach for constrained continuous optimization" in Proceedings of the Eighth International Conference on Contemporary Computing (IC3), (Noida, India), ISBN 978-1-4673-7948-9, vol. 1, pp. 13-18.
- Manuel L., Cedric C., Nadia E.M., Aida O. 2016. "Gene Tree and Correction using Super Tree and Reconciliation". *IEEE Transactions on Computation Biology and bioinformatics* vol. PP, pp. 1-1.
- Nguyen Su, Zhang Mengjie, Johnston Mark, 2014. "A sequential genetic programming method to learn forward construction heuristics for order acceptance and scheduling" in Proceedings of the 2014 IEEE Congress on Evolutionary Computation (CEC), (Beijing, China), ISBN 978-1-4799-1488-3, vol. 1, pp. 1824-1831.
- Ogden H, Rosenberg MS. Multiple sequence alignment accuracy and phylogenetic inference. *Syst Biol.* 2006; vol 55, pp. 314-28.
- Qin A.K., Huang V.L., and Suganthan P.N. 2009. "Differential Evolution Algorithm with Strategy Adaptation for Global Numerical Optimization ". *IEEE Transactions on Evolutionary Computation*, vol. 13, pp. 2827-2839.
- Ranjani R., and Ramyachitra D. 2017. "Application of Genetic Algorithm by Influencing the Crossover Parameters for Multiple Sequence Alignment". 4th IEEE International Conference on Electrical, Computer and Electronics (UPCON), Mathura, India, pp. 33-38.
- Sri R. Leena, Balaji N. 2013. "Meta-heuristic hybrid dynamic task scheduling in Heterogeneous computing environment" in Proceedings of the 2013 International Conference on Computer Communication and Informatics (ICCCI), (Coimbatore, India), ISBN 978-1-4673-2907-1, vol. 1, pp. 1-6.
- Sugawara Kohei, Fujita Hamido, 2013. "A workflow optimization by handling interruption with meta-heuristic approach" in Proceedings of the 12th IEEE International Conference on Intelligent Software Methodologies, Tools and Techniques (SoMeT), (Budapest, Hungary), ISBN 978-1-4799-0421-1, vol. 1, pp. 243-246.
- Sugawara Kohei, Fujita Hamido, 2014. "A workflow optimization by handling subjective attributes with meta-heuristic approach" in Proceedings of the 10th International Conference on Natural Computation (ICNC), (Xiamen, China), ISBN 978-1-4799-5151-2, vol. 1, pp. 497-502.
- Thompson J. D., P. Koehl, R. Ripp, and O. Poch, 2005. "BALiBASE 3.0: Latest developments of the multiple sequence alignment benchmark," *Proteins*, vol. 61, no. 1, pp. 127-136.
- Usman R. and Dennis R.L. 2005. "Probalign: Multiple sequence alignment using partition function posterior probabilities". *Bioinformatics*, vol. 22, pp. 1-16.
- Wei F.G., Ling L.H., San Y.L. and Cai D. 2015. "Artificial Bee Colony Algorithm Based on Information Learning Gene Tree and Correction using Super Tree and Reconciliation ". *IEEE Transactions on cybernetics* vol. 45, pp. 2827 - 2839.
- Yue J. G. *et al.*, 2016. "Genetic learning particle swarm optimization," *IEEE Transaction on Cybernetics*. vol. 46, pp. 2277-2290.