Using Bioinformatics Application to Predict and Detect Single Nucleotide Polymorphism (SNP) in Eukaryotic Organism

Hala F. Hassan[1], Hisham F. Mohammad[2]

1Department of Marine Biology, Marin Science Centre, University of Basrah, Basrah, Iraq
2Department of Biology, College of Science, University of Basrah, Basrah, Iraq

*Corresponding Author E-mail: hala.hassan@uobasrah.edu.iq

Received 07/01/2026 Accepted 14/03/2026 Published 25/03/2026

Abstract

The rapid development of bioinformatics has made data analysis essential for understanding biological processes requiring advanced molecular techniques for detection and prediction based on specific artificial intelligent program and algorithm such as ANN (Artificial Neural Network), The main objective of this research to predictive for rotifers (*Brachionus plicatilis* and *Brachionus calyciflorus*) depending on genetic data which including fifty samples from rotifers species collected from different area in Basrah city water. Fifty samples of the ***18S rRNA*** gene products were subjected to PCR amplification and sequencer also applied artificial neural network program to predict and detect SNP, this study considered one of the very encouragement research for highly confirmed technique in different types of rotifers with short time and low cost application.

Keywords: ANN, Artificial intelligence, ***18S rRNA*** gene, Nucleotide Rotifers.

Introduction

The phylum Rotifera is a relatively small group of microscopic aquatic or semi-aquatic invertebrates, encompassing about 2,000 species of unsegmented, bilaterally symmetrical pseudocoelomatest (Brown *et al.*, 2020). They increase in marine fish larvae rearing around the world has been due partly to the availability of rotifers. These organisms are valuable live food for fish larvae and crustacean culture (Slathia *et al.*, 2024). Several characteristics of rotifers, including their nutritional quality, body size and relatively slow motility have contributed to their usefulness as good prey for active larvae, nature, rapid production rate and the possibility of artificially manipulating its nutritional qualities along with the euryhaline nature (Faudzi *et al.*, 2024).

P-ISSN: 1812-237X, E-ISSN: 2788-5720, https://ijaqua.uobasrah.edu.iq/index.php/jaqua/index

The species under study, the monogonont rotifer *Brachionus plicatilis*, and *Brachionus calyciflorus* is a zooplanktonic invertebrate, typically found in water bodies where environmental factors restrict population growth to short periods lasting months. Then adverse conditions for growth include evaporation of water in temporary habitats leading to desiccation, unfavorable temperatures and lack of food or occurence of predators (Becker *et al*., 2015). This is a relatively hostile environment and the survival of the population in such conditions is ensured via the production of resting eggs. These show a remarkable tolerance to unfavorable conditions and may be stored for decades (Laakmann *et al*., 2025).

Traditionally, rotifers monitoring has been conducted using zooplankton net samples and microscopic identification based on morphological features (Boersma *et al*., 2015). Additionally, molecular approaches allow for a faster taxonomic processing of net samples and readily available data in bioinformatics format, Many studies have been focused on the taxonomy of rotifer in different water bodies (Abd Al-Rezzaq, 2014). The 1*8S rRNA* gene is the structural RNA for the small component of eukaryotic cytoplasmic ribosomes, and thus one of the basic components of all eukaryotic cells. It is the eukaryotic nuclear homologue of 16S ribosomal DNA in Prokaryotes. Sequence data from the gene *18S rRNA* is widely used in molecular analysis to reconstruct the evolutionary history of organisms, especially in vertebrates, as its slow evolutionary rate makes it suitable to reconstruct ancient divergences (Qin *et al*., 2023).

In the last decade, as a result of the widespread use of PCR and DNA sequencing of *18S rRNA* has played a pivotal role in the accurate identification of *Brachionus* isolates and in the discovery of new species. The *18S rRNA* gene is one of the most frequently used genes in phylogenetic studies and an important marker for random target polymerase chain reaction (PCR) in environmental biodiversity screening. In general, rRNA gene sequences are easy to access due to highly conserved flanking regions allowing for the use of universal primers. Their repetitive arrangement within the genome provides excessive amounts of template DNA for PCR, even in smallest organisms (Zheng *et al*., 2022). Bioinformatics is an interdisciplinary field that combines biology, computer science, and information technology to analyze and interpret biological data, especially genomic information (Adie, *et al*., 2025).

It involves the development of algorithms, databases, and software tools to understand biological processes at the molecular level through computational methods. Also it plays a crucial role in genomics, and other-omics area , in addition to helping in understanding biological processes and predicting gene functions (Satya *et al*., 2024). Artificial Neural Network (ANN) is one of the greatest suppliers in the field and genomic algorithms greatly benefit from the increased availability of data due to the big size and complexity of this data requires of computational methods. On the other hand, these methods also face challenges such as accuracy and performance which are key priorities in the development of algorithms matching and improving the data analysis (Iqbal and Kumar, 2023). The power of machine learning to predict the existence and function of genetic structures is restricted by data collection (Gupta *et al*., 2025).

Materials and Methods

Samples collection

Fifty samples were collected during December of 2023 to March of 2024 from eight selected stations in Shatt Al-Arab River by conical net of 1meter length and 40 cm diameter with mesh size of 50 µm. Its thrown into the water and pull to a distance of 3 meters by tied rope and then filtrate were poured in plastic sterile containers and transferred to laboratory (Hammadi, 2010).

The purification of samples

The samples were purified immediately after reached the laboratory for purity after phynotypic diagnosis by using in dissecting microscope according to taxonomy key (Fontaneto, 2010; Hammadi *et al*., 2012). To isolate the rotifers, two sieve [90 and 43 micron] were used , the first one used to obscure the large organisms and the second one (43 micron) were used to collect the rotifers which assemble on the top sieve by washing the sieve with sterilized water, then anatomical microscope used to exam the rotifers and collect the *Brachionus* species only according to the key and put in plastic tanks (Hassan *et al*., 2015).

Molecular methods

DNA Extraction

The extraction of DNA done from the fresh sample (Brachionus species) according to the protocol of kit by PROMEGA (Serial No: A7710).

Detection of genomic DNA by agarose gel electrophoresis

Electrophoresis was performed using four reagents TBE buffer, Bromo phenol blue, Ethidium bromide and Agarose powder according to Sambrook and Russel (2001) and Hassan *et al.* (2015). The procedure start with 25 ml of 1× TBE buffer when mixed with agarose powder (0.2g) then heated in a microwave for 30 seconds until the solution became clear and cool to (50-60 ^{0}C) then added 0.2 µl of ethidium bromide (10mg/ml). The tray was attached to the gel, and the comb was positioned at one side of the tray. The agarose mixture was poured into the tray and allowed it to solidify at room temperature. The comb was gently removed and 1× TBE was added until the buffer covered over the surface of the gel. Approximately 6 µl of DNA was mixed with 3µl of bromophenol blue and loaded in the wells of 0.8% agarose gel after that cathode linked to the wells sides and anode to the opposite side. Then, the power turned on 60V until bromophenol blue tracking dye moved to other side. DNA was detected and analyzed under UV transilluminator.

Identification of Brachionus sp. by using specific *18S rRNA* amplification

The *18S rRNA* gene was amplified using primers corresponding to conserved regions as 200bp (Winnepenninckx *et al.*,1995).

Table (1): Oligonucleotide primer sequences used for PCR amplification of *18S rRNA* gene.

Primers	**Sequence**	**Length**
Forword	5-AGATTAAGCCATGCATGCGTAAG-3	23bp
Reverse	5-TGATCCTTCTGCAGGTTCACCTAC-3	24bp

Table (2): Reagents of PCR amplification (50 µl) for *18S rRNA*.

No	**Reagent**	**Volum**
1	DNA	10µl
2	Forward Primer	2µl
3	Reverse Primer	2µl
4	Master Mix .2x	11µl
5	Nuclease-free water	25µl
	Total	50µl

The Touch down PCR program for Specific *18S rRNA* gene amplification consisted of an initial denaturation step at 95°C for 2 minutes, followed by 30 cycles of denaturation at 95°C for 30 seconds, annealing at 61.3°C decrease 0.5°C percycle foe 30 seconds, and extension at 72°C for 20 seconds, then repeat steps 2 to 4 for 14 more time., A final denaturation 95 °C for 30 seconds, followed the 54.3 °C for 30 seconds, In addition to 72 °C for 20 seconds, repeate steps 6 to 8 for 19 time. While the final extentsion step was performed at 72 °C for 5 minutes.

Identification of Brachionus species

The sequence (Query) after treatment and correction of the sequencer product was copied and pasted in the box of the "BLAST" from the website http://www.blast.ncbi.nlm.nih.gov. which is followed by the National Center for Biotechnology Information (NCBI), then the name of the *Brachionus* species will appear directly with the percentage of its compared sequence (Subject). Only the similarities of >or = 99% were depended for the diagnosis in the present study.

The Analysis of the sequence data

All the corrected sequences of *18S rRNA* data from the present study, by http://www.blast.ncbi.nlm.nih.gov (Becker *et al.*, 2004). were aligned for the concatenated of 200bp (according to the shorter sequence) at the same sequence ends by cutting, then pasted together in the box of "CLUSTALW" http://www.ebi.ac.uk/clustalw/(Thompson *et al.*, 1994). This program performs multiple alignments on nucleotide sequences, this step was performed prior to tree construction using MAFFT and T COFFE(Multiple alignment program for amino acid or nucleotide sequences) Version8 http://mafft.cbrc.jp/alignment/server/. A multiple alignment is built up heuristically by a series of pair wise alignments, the most closely related sequences were aligned first and the more distant ones were added. MAFFT uses the Neighbour Joining (NJ) method for tree construction (Katoh *et al.*, 2002). In addition to producing a multiple sequence alignment, the MAFFT website also provides a tree file, with a DNA extension, which can be opened by forester-1027 (Zmasek and Eddy, 2001).

Results

Identification of Brachionus species by*18S rRNA*

The extracted DNA from each isolate (n=50) was subjected to PCR for amplifying *18S rRNA*, So the individual band of the gene was characterized by 200 bp due to comparison with the standard molecular DNA Ladder (100-1000bp) as shown in Figure (1).

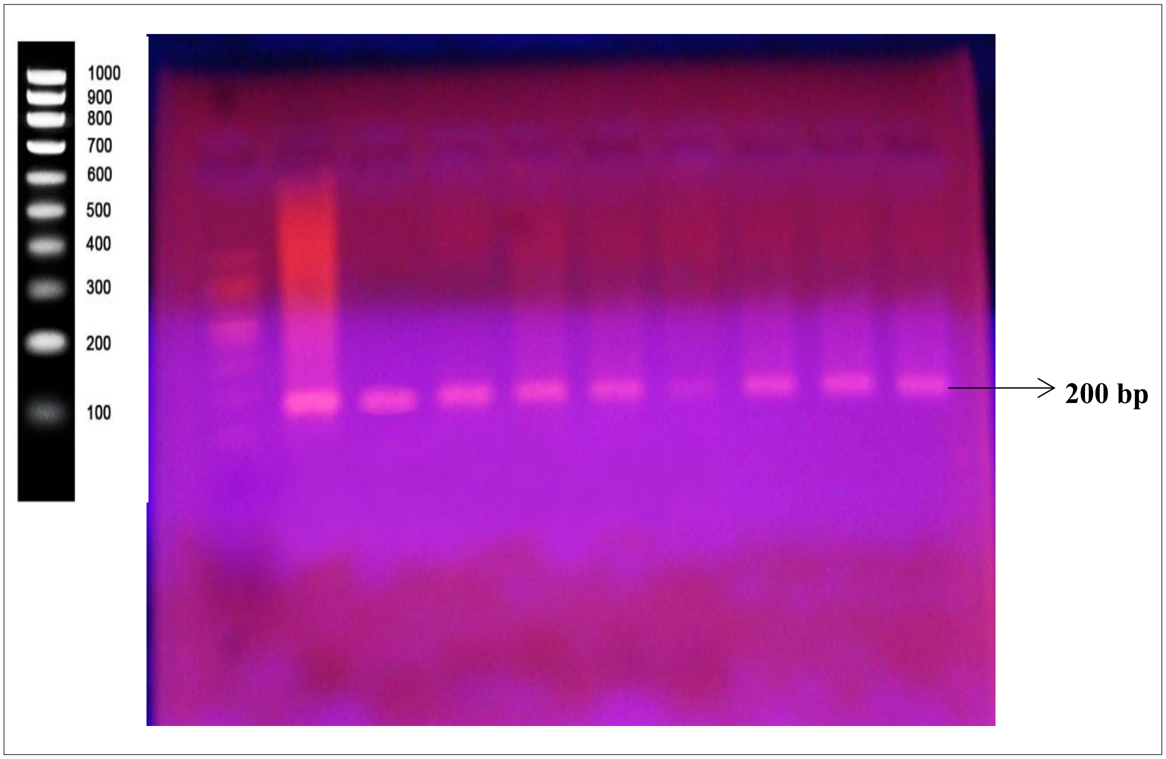


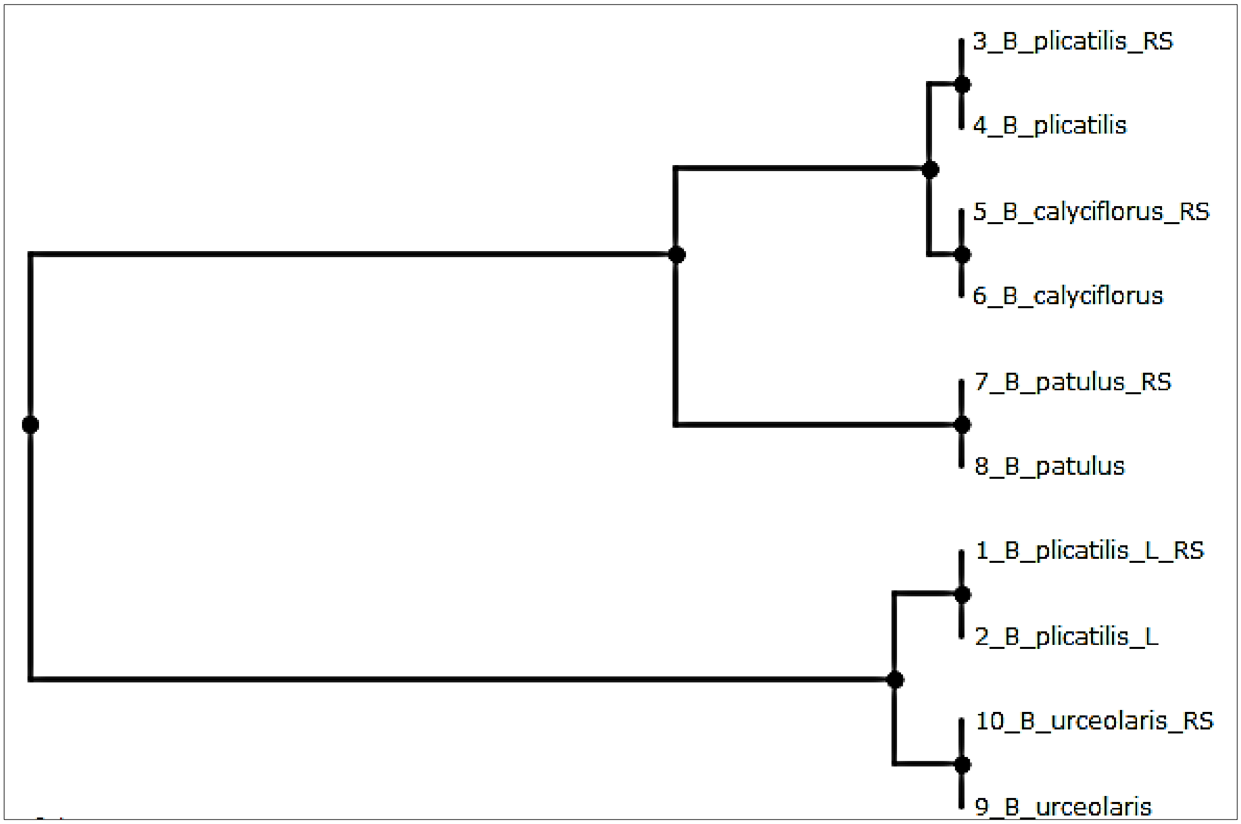
Figure (1): Agarose gel (2%gm and 60V,2MA) electrophoresis patterns show PCR amplified products of gene *18S rRNA*. Lane L: 100-1000bp ladder, lanes 1- 9: Gene *18S rRNA* bands of Brachionus sp.

Sequencing for *18S rRNA* and identification of Brachionus species.

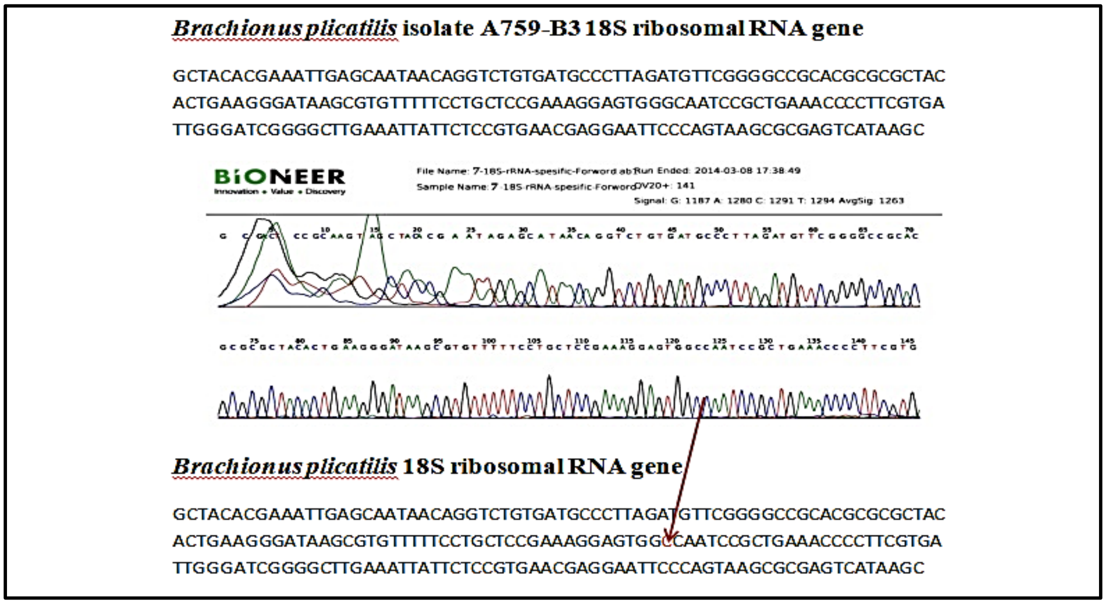
Our results for the identification of *18S rRNA* nucleotides for fifty samples of Brachionus isolates were collected from various locations in Al-Basrah city. During the current study, 27 *Brachionus* samples were identified by sequence included *Brachionus plicatilis* was the common species (22%), followed *B. calyciflorus* (20%), while *B. urceolaris* and *B.patulus* were (6%) for each one, while the other samples from their were not identified *Brachionus* were (46%) (n=23) because of the sequencing of these isolates failed to give results.

Phylogenetic tree of Brachionus species

The Nucleotide sequences data of all Brachionus species were concatenated producing a sequence length (200) bases depended on all species. The rooted phylogenetic tree was constructed and visualized in Figure (2) for Brachionus species. Individual NJ tree were drawn for each .All data which share the same sequences in number were pasted to gather in the "Clustal W" and "TREE RVIEW" was employed to produce the tree. These trees show the distribution and phylogenetic relationships among the studied Brachionus species and their identical reference strains.



Figure(2): Rooted Neighbour Joining phylogenetic tree constructed from concatenated sequences of 200 bp for each isolate (derived from an alignment of *18S rRNA* sequences) then produced from a MAFFT alignment version 8 and visualized using forester version1027.



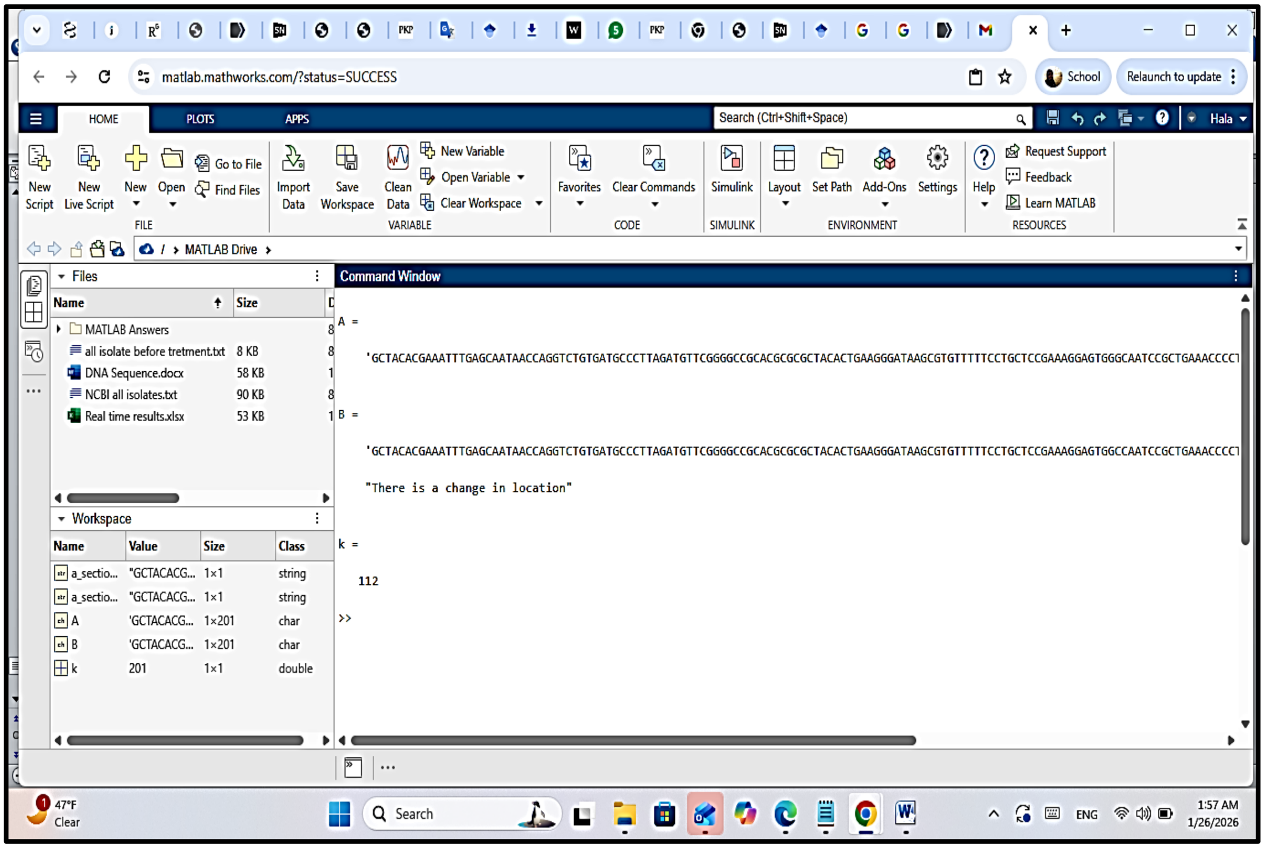
Brachionus plicatilis **isolate A759-B3 18S ribosomal RNA gene**

GCTACACGAAATTGAGCAATAACAGGTCTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTAC
ACTGAAGGGATAAGCGTGTTTTTCCTGCTCCGAAAGGAGTGGGCAATCCGCTGAAACCCCTTCGTGA
TTGGGATCGGGGCTTGAAATTATTCTCCGTGAACGAGGAATTCCCAGTAAGCGCGAGTCATAAGC

Brachionus plicatilis **18S ribosomal RNA gene**

GCTACACGAAATTGAGCAATAACAGGTCTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTAC
ACTGAAGGGATAAGCGTGTTTTTCCTGCTCCGAAAGGAGTGGCCAATCCGCTGAAACCCCTTCGTGA
TTGGGATCGGGGCTTGAAATTATTCTCCGTGAACGAGGAATTCCCAGTAAGCGCGAGTCATAAGC

Figure (3) Comparison of *18S rRNA* nucleotide sequences (200bp) for the isolate *Brachionus plicatilis* (with peaks) from present study and reference isolate A759-B3. A Point mutation type Transversion (C instead G) at the position 112bp changing the amino acid Alanine to Proline.



Name	Size
MATLAB Answers	
all isolate before tretment.txt	8 KB
DNA Sequence.docx	58 KB
NCBI all isolates.txt	90 KB
Real time results.xlsx	53 KB

Name	Value	Size	Class
a_sectio...	"GCTACACG...	1×1	string
a_sectio...	"GCTACACG...	1×1	string
A	'GCTACACG...	1×201	char
B	'GCTACACG...	1×201	char
k	201	1×1	double

```
A =

    'GCTACACGAAATTTGAGCAATAACCAGGTCTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTACACTGAAGGGATAAGCGTGTTTTTCCTGCTCCGAAAGGAGTGGGCAATCCGCTGAAACCCCT

B =

    'GCTACACGAAATTTGAGCAATAACCAGGTCTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTACACTGAAGGGATAAGCGTGTTTTTCCTGCTCCGAAAGGAGTGGCCAATCCGCTGAAACCCCT

    "There is a change in location"

k =

   112

>>
```

Figure (4): *Brachionus plicatilis* Point mutation type Transversion (C instead G) at the position 112bp by ANN mutation detection at 643 nucleotides position.

Discussion

The results shed light of using artificial neural network program (ANN) to detect mutation According to Lu, *et al.*, (2014), and Naito (2019), they clearfield the available computational methods, such as machine learning has been applied widely and achieved promising results. The training and testing data accuracies of the ANN model was 100%. For (ANN) were used to train different prediction models based on molecular identification (Charan *et al.*, 2024). Primers were used in the present study to amplify the *18S rRNA* gene for all of the DNA of Brachionus which were collected from the area of study. All identified Brachionus sp. (50) isolates including (27) for*18SrRNA*. While other isolates were unidentified. The present study approves the results of the previous studies which showed the same domains of these species like (Papakostas *et al.*, 2006). The present study results are in agreement with the results of (Winnepenninckx *et al.*, 1995). Therefore, the aim of the present study is to detect the genetic differences of the isolates of Brachionus sp. (Ray, 2022).

Conclusion

Our implemented a new method using machine learning to detect mutation using sequence-based Artificial Neural Networks. According to the results were more important for detect mutation in different types of living things. Hence, the observations were integrated by ANN leads to moderate performance in detecting a mutation that causes amino acid change. Mutation refers to a change in the DNA sequence of a gene or chromosome in an organism leading to the development of a new characteristic or trait absent in the parents type.

References

Abd Al-Rezzaq, A.J. (2014). A Dignostic and Ecological Study of the Planktonic Species of Rotifera in the Al-Hilla river-Iraq. PhD.thesis, Babylon Univ, College of Sci.(in Arabic).

Adie, A. E.; Beshel, J. A.; Eze, V. H. U.; Bubu, P. E.; Abreka, M.; Maduabuchi, E. C., and Igwe, M. C. (2025). Bioinformatics and Genomics; the Integration of Computational Tools in Understanding Biological Data. JMSAE, Jour Mater Sci & Appl Eng. 4 (4): 1-7.

Becker, K.; Harmsen, D.; Mellmann, A.; Meier, C.; Schumann, P., and Peters, G. (2004). Development and evaluation of a quality-controlled ribosomal sequence database for 16S ribosomal DNA-based identification of Staphylococcus species. J. Clin. Microbiol. 42: 4988- 4995.

Becker, M.; Wiltshire, K. H.; Kong, S. M.; Greve, W., and Renz, J. (2015) Long-term change in the copepod community in the southern German bight. J. Sea Res., 101, 41–50.

Brown, P. D.; Schröder, T.; Ríos-Arana, J. V.; Rico-Martinez, R.; Silva-Briano, M.; Wallace, R. L., and Walsh, E. J. (2020). Patterns of Rotifer Diversity in the Chihuahuan Desert. Diver, 12(10), 393.

Charan, E. S.; Sharma, A.; Sandhu, H., and Garg, P. (2024). FGFR1Pred: an artificial intelligence-based model for predicting fibroblast growth factor receptor 1 inhibitor. Mol. Divers, 28(4): 065-2076.

Faudzi, N. M.; Shapawi, R., and Fui, C. F. (2024). The Importance of Rotifer as Live Feed. Essentials of Aqua. Practices, 41.

Fontaneto, D.(2010). Rotifera Bdelloidea. Summer school in taxonomy ,Valdieri, Itally ,11pp.

Gupta, S.; Janu, N.; Nawal, M., and Goswami, A. (2025). Genomics and Machine Learning: ML Approaches, Future Directions and Challenges in Genomics. Genomics at the Nexus of AI, Computer Vision, and Machine Learning, 437-457.

Hammadi, N.S.; Salman, D.S., and Al-Essa, S.A. (2012). Rotifera of Shatt Al-Arab Region Iraq. Basrah University, Publication about Marin Sci.Center, 258 p.

Hammadi, N.S . (2010) An Ecological Study of the Rotifera of Shatt Al-Arab Region. PhD. thesis , College of Agric. Univ. of Basrah.

Hassan, H.F; Al-Badran, I.A., and Ali, M.H. (2015). “Morphological and Molecular Identification of Collected Brachionus spp From Shatt Al-Arab River”. MS thesis, Univ. of Basrah, Basrah, Iraq.

Iqbal, N., and Kumar, P. (2023). From Data Science to Bioscience: Emerging era of bioinformatics applications, tools and challenges. Procedia Comp Sci, 218, 1516-1528.

Katoh, K.; Misawa, K.; Kuma, K., and Myata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucliec Acids Research. 30(14):3059-3066.

Laakmann, S.; Cornils, A.; Metfies, K.; Koplin, J.; Neuhaus, S.; Bunse, C., and Flores, H. (2025). Of sequences and images-diversity and quantity of Arctic epipelagic zooplankton by an integrative approach. J. Plankton Res., 47(6): fbaf059.

Lu, Y.; Deng, J.B.; Carson, M.; Lu, H., and Lu, L. (2014). Computational methods for the prediction of microbial essential genes. Curr. Bioinform, 9(2): 89-101.

Naito, T. (2019). Predicting the impact of single nucleotide variants on splicing via sequence-based deep neural networks and genomic features. Human mutation, 40(9): 1261-1269.

Papakostas, S.; Dooms, S.; Triantafyllidis, A.; Deloof, D.; Kappas, I.; Dierckens, K.; De Wolf, T.; Bossier, P.; Vadstein, O.; Kui, S.; Sorgeloos, P., and Abatzopoulos, T.J. (2006). Evaluation of DNA methodologies in identifying Brachionus species used in European hatcheries. Aquac. 255, 557–564.

Qin, W.; Wang, S.; Xia, P.; Tang, F., and Zhao, Y. (2023). Molecular characterization and phylogenetic analysis of Paratrichodina africana Kazubski and El-Tantawy, 1986 based on 18S rRNA gene data with the evolutionary hypothesis of trichodinids. Parasitol. Int., 94, 102735.

Ray, A. (2022). DNA mutation, repair, and recombination. In Genetics Fundamentals Notes (433-490).Springer Nature Singapore.

Sambrook, J. and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, USA. (1):112–1.118.

Satya, S.; Sharma, S.; Choudhary, G., and Kaushik, G. (2024). Advances in Environmental Microbiology: A Multi-omic Perspective. In Microbial Omics in Environment and Health (175-204).

Slathia, D.; Kour, S., and Kour, S. (2024). First report of Macrochaetus sericus Thorpe, 1893 and Lecane tenuiseta Harring, 1914 (Rotifera: Monogononta) from Jammu waters (J&K), India. J. Threat. Taxa, 16(3): 24923-24929.

Srivastava, U.; Kanchan, S.; Kesheri, M.; Gupta, M.K., and Singh, S. (2024). Types of omics data: genomics, metagenomics, epigenomics, transcriptomics, proteomics, metabolomics, and phenomics. In Integrative Omics (13-34). Academic Press.

Thompson, J. D.; Higgins, D. G., and Gibson, T. J. (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penlties and weight mutrix choice. Nucleic Acids Res. 22: 4673-4680.

Winnepenninckx, B.; Backeljau, T.; Mackey, L. Y.; Brooks, J. M.; De Wachter, R.; Kumar, S., and Garey, J.R. (1995). 18S rRNA data indicate that Aschelminthes are polyphyletic in origin and consist of at least three distinct clades. Molecular Biology and Evolution, 12(6): 1132-1137.

Zheng, X.; He, Z.; Wang, C.; Yan, Q., and Shu, L. (2022). Evaluation of different primers of the 18S rRNA gene to profile amoeba communities in environmental samples. Water Biology and Security, 1(3): 100057.

Zmasek, C. M., and Eddy, S. R. (2001). ATV: display and manipulation of annotated phylogenetic trees. Bioinform. 17: 383-384.

استخدام تطبيقات المعلوماتية الحيوية للتنبؤ بتعدد أشكال النوكليوتيدات المفردة (SNP) والكشف عنها في الكائنات حقيقية النواة

حلا فاضل حسن[1] و **هشام فياض محمد**[2]

1 قسم الأحياء البحرية، مركز علوم البحار، جامعة البصرة
2 قسم علوم الحياة، كلية العلوم، جامعة البصرة

*Corresponding Author E-mail: hala.hassan@uobasrah.edu.iq

تاريخ الإستلام: 2026/01/07 تاريخ القبول: 2026/03/14 تاريخ النشر : 2026/03/25

المستخلص

أدى التطور السريع لعلم المعلوماتية الحيوية إلى جعل تحليل البيانات ضرورياً لفهم العمليات البيولوجية، مما يتطلب تقنيات جزيئية (ANN) متقدمة يهدف للكشف والتنبؤ بالاعتماد على برامج وخوارزميات ذكاء اصطناعي محددة، مثل الشبكات العصبية الاصطناعية يعتمد هذا البحث بشكل أساسي إلى التنبؤ بالروتيفرا (*Brachionus plicatilis*) و (*Brachionus calyciflorus*) بناءً على البيانات الجينية، وذلك من خلال دراسة خمسين عينة من أنواع مختلفة من الروتيفرا جُمعت من مناطق مختلفة في مياه مدينة البصرة. خضعت **18S rRNA** نتائج الجين لخمسين عينة لتضخيم PCR وتسلسلها، كما طُبّق برنامج للتنبؤ والكشف عن تعدد أشكال النوكليوتيدة، وتعد الدراسة من الأبحاث الواعدة التي تُقدم تقنية عالية الدقة في أنواع مختلفة من الروتيفرات، مع تطبيق سريع ومنخفض التكلفة.

الكلمات المفتاحية: الشبكات العصبية الاصطناعية، الذكاء الاصطناعي، تسلسل النوكليوتيدات الروتيفرات جين (**18S rRNA**).