Archives of Razi Institute, Vol. 76, No. 4 (2021) 895-901

DOI: 10.22092/ari.2021.355942.1743

Original Article

Analytical Computation of Interleukin17A Activity in Breast Cancer Patients Using Bioinformatics Methods

Mohammed, A. S[1 *], Al-Janabi, A. A[1]

1. Department of Applied Sciences, University of Technology, Baghdad, Iraq

Received 22 August 2021; Accepted 19 September 2021
Corresponding Author: bannsaeed@yahoo.com

Abstract

Interleukin-17A (IL-17A) is a member of the Interleukin-17 family, which belongs to the pro-inflammatory cystine-knot cytokines. Recent studies on the etiology of breast cancer have focused on the role of immunity and inflammation. The pro-inflammatory cytokines IL-17A can medicate cancer-related inflammation. The present study aimed to analyze the mutation in physicochemical properties and structure of the *Interleukin-17A* gene in developing breast cancer using bioinformatics methods. A total of 60 blood samples were obtained from Iraqi women aged 25 to 75 with breast cancer. Twenty blood samples were obtained from healthy women in the same age range as a control group. Deletion and missense mutations detected by BLAST in samples with breast cancer. The present study determined the physicochemical properties of IL-17A such as hydrophilic nature, alpha-helical and 3D structure. The results of this study indicated that IL-17A is considered a marker for a patient with breast cancer. Also, mutations in the *IL-17A* gene affect the structure and physicochemical properties of the IL-17A protein complex.

Keywords: Bioinformatics, Breast Cancer, IL-17A

1. Introduction

Bioinformatics involves combining various scientific disciplines such as applied mathematics, statistics, and computer science and using the latest methods for solving biological problems (1). Nowadays, bioinformatics applications are becoming more crucial in academic studies, which analyze the structures of amino acids and proteins. These studies design metabolic pathways for division and prediction of function as well as localization of their position in cells using alignment operations (2). Various bioinformatics-related applications are available that should be used for conducting new experiments and researches due to their roles in simply organizing and analyzing data to access and interpret (3). The human *IL-17A* gene has 1874 base pairs in the short arm of chromosome 6 region 1 band 2 sub-band 2(Chr. 6 p12.2) which has 3exons, 2introns and is cloned from CD4+ T cells. Each member of the IL-17 family has a unique cellular expression pattern. IL-17A and IL-17F expression appear to be limited to a small number of activated T cells and increase during inflammation (4). IL-17A is a 155-amino acid protein that is a disulfide-linked, homodimeric, secreted glycoprotein with a molecular mass of 35 kDa. Each subunit of the homodimer is approximately 15-20 KDa. IL-17has a 23-amino-acid signal peptide followed by a 123-amino-acid chain region that is typical of the IL-17 family. Purification of the protein revealed two bands at 15 KDa and 20 KDa, which discovered an N-linked glycosylation site. Four conserved cysteines that produce two disulfide linkages were discovered by comparing different members of the IL-17 family. IL-17 differs from other interleukins. Furthermore, no proteins or structural domains are

found similar to IL-17 (5). Breast cancer is a subset of diseases in which cells of breast tissue keeps uncontrollably change and divide, typically forming a lump or mass. Various types of breast cancers mostly start in the lobules (milk glands) or ducts that connect the lobules to the nipple (6). The present study aimed to analyze and detect the sequence, compute, and predict the structure of IL-17A in patients with breast cancer using a bioinformatics tool.

2. Material and Methods

2.1. Blood Sample Collection

A total of 60 blood samples were obtained from Iraqi women aged 25 to 75 with breast cancer. Twenty blood samples were obtained from healthy women in the same age range as a control group. Blood samples were collected from each patient in healthy and control groups with venous blood sampling using disposable syringes in sterile EDTA tubes. Approximately 3 ml of blood samples were stored from each patient in healthy control the blood samples at –20°C (7).

2.2. DNA Extraction

DNA extraction is a method to purify DNA from cell debris by disrupting the cell membrane. Easy Pure® Blood Genomic DNA Kit was used to extract DNA from the blood of participants in both groups.

2.3. Primer Design for *IL-17A* Gene

The *IL-17A* gene sequence was collected from the Genome Database of the National Center for Biotechnology Information (NCBI) (RefSeq: NC_000006.12). Primer3 software was used (http://bioinfo.ul.ee/primer3) for rs2275913 of IL-17A primer design (8, 9). The sequences of the forward and the reverse primers were 5'-GGCCAAGGAATCTGTGAGGA -3' and 5'-GGGATGGATGAGTTTGTGCC -3', respectively. The product size of rs2275913 was 441bp, then the primers provided by Alpha DNA company /Canada were used as lyophilized for the primers in PCR.

2.4. Polymerase Chain Reaction (PCR)

The IL-17A primer was used to amplify the *IL-17A* gene by PCR according to the instruction of EasyTaq PCR SuperMix. The mixture of PCR solution is shown in table 1 and table 2.

Table 1. The mixture of PCR solution

Component	Volume
Master mix EasyTaq® PCR SuperMix	12.5µl
Forward primer	1 µl
Reverse primer	1 µl
DNA	4µl
Nuclease free water (N.F.W)	7.5µl

Table 2. PCR conditions designed for the study

Stage	Function	Temperature	Time	Cycle
Stage 1	Denaturation	95ºC	30 s	1
Stage 2	Denaturation	95ºC	5 s	35
	Annealing	58ºC	30 s	
	Extension	72ºC	30 s	
Stage 3	Extension	72ºC	60 s	1

2.5. Preparation of 1% Agarose Gel

Agarose gel electrophoresis was used to confirm the presence of amplification after PCR amplification or DNA extraction. PCR was completely dependent on the criteria for extracted DNA. The solutions used in this study included 1X TBE buffer, loading dye, DNA ladder marker healthy, and Gel stein (ethidium bromide). DNA samples of 60 cancer patients and 20 healthy controls were sent to Korea for sequencing at Macrogen Corporartion Company by using an automated DNA sequencer to confirm the PCR products of the *IL-17A* gene. DNA samples from 60 cancer patients and 20 healthy controls were sent to Korea for sequencing at Macrogen Corporartion Company using an automated DNA sequencer. Then, bioinformatics methods were used to analyze the function and predict the structures of IL-17Aprotein.The BLASTX program was used to translate nucleotide sequence to that of the amino acid (protein sequence), the results of which are used to detect mutations in the rs2275913 region in the *IL-17A* gene sequence, include substitution, frameshift, missense, and deletion mutations in breast cancer patients (10). The physiochemical properties of the primary protein composition were calculated by comparing the results with those of the control samples using ProtParam program which showed the effect of mutations on the molecular weight and protein stability (11).

3. Results and Discussion

3.1. Detection of *IL-17A* Gene by PCR Technique

As shown in figure 1, the results for the amplification segments in the DNA of patients and control of the first primer gave a clear band with a size of 441 bp when electrophoresed on (1%) agarose gel at 100 volts for 60 min. Bands were shown in UV light after staining with a red safe stain. Lane M: DNA ladder (in base pairs) from bottom to top: 100, 200, 300, 400, 500, 600, 800, 1000, 1200,1300, 1400 and 1500.

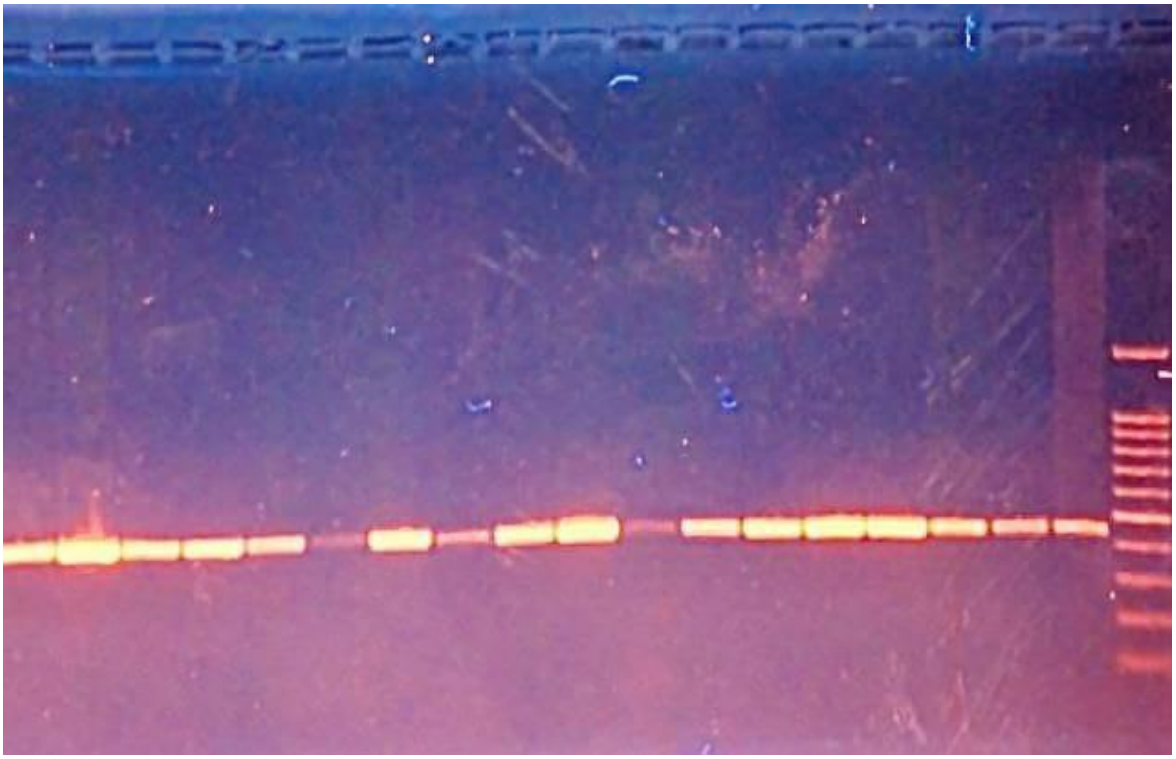

Figure 1. PCR analysis for rs2275913 of *IL-17A* gene in the extracted DNA, with band size (441 bp) using the first primer

3.2. Sequence Alignment of *IL-17A* Gene

Several mutations exist in the gene that replicate in some samples of the same type and locations, indicating that some mutations may be repetitive. Table 3 presents the frequency of these mutations.

The results of the primer showed that one deletion mutation and multiple substitution transition appeared in all samples with breast cancer as following: at locations 18, 65,200, and 246 guanine was substituted by adenine (G → A), at locations 22, 32, and 55cytosine was substituted by thymine (C → T), at location 26 a deletion was found in adenine(A → -),at location 30 adenine was substituted by thymine (A → T), at location 47 adenine was substituted by guanine (A → G), at location (41,63) thymine was substituted by adenine (T → A), at location 213 guanine was substituted by cytosine (G → C), at location 161 thymine was substituted by guanine (T → G), and at location 379 thymine was substituted by cytosine (T → C). The samples of the healthy control group showed transition and substitution of the breast patients as follows: at location (22, 32) cytosine was substituted by thymine (C → T), at location 30 adenine was substituted by thymine (A → T), at location 47 adenine was substituted by guanine (A → G), at location (41,63) thymine was substituted by adenine (T → A), at location 246 guanine was substituted by adenine (G → A).The results indicated that mutations significantly affect the results. The translation explains that a missense mutation leads to amino acid substitution. Deletion causes the frameshift mutation. Table 4 shows mutations that affect the protein translation process in breast cancer. The mutations of the *IL-17A* gene in patients with breast cancer were recorded in the present study. These mutations affect the structure of the protein in comparison with IL-17 A retrieved from NCBI, as shown by structural analyzes such as the numbers and positions of alpha-helix, β- turn, and coil. This result is consistent with those of studies (12, 13) which state that mutation significantly changes the sequences and affects the structure of the protein. Also, the present study showed that protein folding pathways are not significantly affected by changes in sequence.

Nucleotides sequences of the *IL-17A* gene were translated to amino acids sequence (protein sequence) for each group of breast cancer patients as well as the healthy controls using BioEdit sequence alignment editor. While the amino acid sequence of the normal IL-17A protein was retrieved from the NCBI database of which consists of 126 amino acids. ProtParam program was used for analyzing the primary structure of the protein which provided the physiochemical properties of the IL-17A protein for healthy control as well as for cancer groups. These properties reflect function, stability, the effect of the protein, and many other features (13). The molecular weight, Isoelectric Point (PI), and Instability Index (II) for all the study groups are listed in table 5.

Table 3. Frequency of some mutations in *IL-17A* gene

Wild Type	Mutant Type	No. of repeats	% of repeats
G	A	22	27.5%
C	T	11	14%
A	T	6	7.5%
A	G	3	3.7%
T	A	63	78%
G	C	60	75%
T	G	13	16%
T	C	5	6%
A	---	80	100%

Table 4. The wild and mutant genetic code of DNA, type of mutation, and the effect of mutation on translation in the proteins of patients with breast cancer in rs2275913 of *IL-17A* gene

Patient number	Wild genetic code of DNA	Mutant genetic code of DNA	Type of mutation at DNA level	Wild genetic code of RNA	Mutations in genetic code of RNA	Effect of the translation
1	AGA	AAA	Transition	AGAArg	AAALys	Missense
2	CCT	CTT	Transition	CCUPro	CUULeu	Missense
3	CAG	C-G	Deletion	CAGGlu	C-G	Frameshift
4	ACC	AAC	Transversion	ACCThr	AACThr	Silent
5	CAG	CTG	Transversion	CAGGln	CUGLeu	Missense
6	CCA	TT-	Deletion	CCAPro	UU-	Frameshift
7	CTG	CGG	Transversion	CUGLeu	CGGArg	Missense
8	CAG	CTG	Transversion	CAGGln	CUGLeu	Missense
9	CTG	CGG	Transversion	CUGLeu	CGGArg	Missense
10	AGA	AAA	Transition	AGAArg	AAALys	Missense
11	TGT	TCT	Transversion	UGUCys	UCUSer	Missense
12	CTG	CGG	Transversion	CUGLeu	CGGArg	Missense
13	CAG	C-C	Deletion	CAGGln	C-C	Frameshift
14	CAG	C-G	Deletion	CAGGlu	C-G	Frameshift
15	CTG	CGG	Transversion	CUGLeu	CGGArg	Missense

Table 5. The physiochemical prosperities ofIL-17A protein was retrieved from the NCBI database using the ProtPram program for the breast cancer patients and healthy control subjects

Group	M.Wt.	PI	Instability Index (II)
Breast Cancer	14821.36	9.61	47.57
	15387.23	9.95	50.30
	15096.30	10.15	51.58
	15849.21	9.93	45.56
	14880.20	9.50	48.91
	15321.48	9.53	53.47
	15621.93	9.95	49.96
	15461.56	9.91	52.97
	15435.23	10.01	47.40
	15185.56	9.73	48.08
Healthy Control	14927.63	9.71	48.68
	15189.12	9.53	49.49
	14905.12	9.71	50.49
	15069.99	9.54	48.07
	14915.88	9.84	51.54
Ref Seq	14687.63	9.91	47.68

The molecular weight of IL-17A protein was decreased or increased for patients with breast cancer compared with that of healthy controls. Instability Index (II) is one of the structure-dependent methods of the primary structure of a protein that is used for in vivo protein stability predictions. All samples were stable but in different ranges. The isoelectric point (PI) of the protein (at which it is least stable) was computed. When the result is less than 7 shows precipitation in acidic buffers, while it is greater than 7 means the solubility is in basic buffers (14, 15). The results of the present study are consistent with the effect of mutations on the primary structure of the protein with the previous experiment (12).

From all the obtained results in this study, it was concluded that the IL-17A protein can be considered as a breast marker to diagnose breast cancer. The effect of mutations on the translation including the stability and the proper protein functions which result in loss or gain of function. Our results are in line with (16, 17) which represent the importance of using bioinformatics to reveal the effect of mutations on the protein structure which affects the stability and function of the produced protein.

Authors' Contribution

Study concept and design: A. S. M.
Acquisition of data: A. A. A.
Analysis and interpretation of data: A. A. A.
Drafting of the manuscript: A. S. M. and A. A. A.
Critical revision of the manuscript for important intellectual content: A. S. M.
Statistical analysis: A. S. M.
Administrative, technical, and material support: A. S. M.

Ethics

All procedures performed in this study involving human participants were in accordance with the ethical standards of the University of Technology, Baghdad, Iraq under the project number of 78541-78514.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Aston KI. Genetic susceptibility to male infertility: news from genome-wide association studies. Andrology. 2014;2(3):315-21.
2. Al-khafaji Z, Al-sheikhly A. Bioinformatics. Baghdad Iraq: University of Al-Nahrain Press; 2012.
3. Hao L, Leng J, Xiao R, Kingsley T, Li X, Tu Z, et al. Bioinformatics analysis of the prognostic value of Tripartite Motif 28 in breast cancer. Oncol Lett. 2017;13(4):2670-8.
4. Hymowitz SG, Filvaroff EH, Yin JP, Lee J, Cai L, Risser P, et al. IL-17s adopt a cystine knot fold: structure and activity of a novel cytokine, IL-17F, and implications for receptor binding. EMBO J. 2001;20(19):5332-41.
5. Kolls JK, Linden A. Interleukin-17 family members and inflammation. Immunity. 2004;21(4):467-76.
6. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. The Lancet. 2011;378(9805):1812-23.
7. Li Q, Kang T, Tian X, Ma Y, Li M, Richards J, et al. Multimeric stability of human C-reactive protein in archived specimens. PLoS One. 2013;8(3):e58094.
8. Koressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics. 2018;34(11):1937-8.
9. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinform. 2012;13:134.
10. Dhary Kamel M, Mohammed A, Ibrahim A. Sequence and Structure Analysis of CRP of Lung and Breast Cancer Using Bioinformatics Tools and Techniques. Biosci Biotechnol Res Asia. 2018;15:163-74.
11. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005. p. 571-607.
12. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary

information. Nucleic Acids Res. 2014;42(Web Server issue):W252-8.

13. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinform. 2010;11:548.
14. Gamage DG, Gunaratne A, Periyannan GR, Russell TG. Applicability of Instability Index for In vitro Protein Stability Prediction. Protein Pept Lett. 2019;26(5):339-47.
15. Tokuriki N, Tawfik DS. Stability effects of mutations and protein evolvability. Curr Opin Struct Biol. 2009;19(5):596-604.
16. Kamimura K. Identification of molecular transition of hepatocellular carcinoma: a novel method to predict the initiation of metastasis. Stem Cell Investig. 2019;6:5.
17. Zhao S, Liu J, Nanga P, Liu Y, Cicek AE, Knoblauch N, et al. Detailed modeling of positive selection improves detection of cancer driver genes. Nat Commun. 2019;10(1):3399.