Benha Journal of Applied Sciences (BJAS) **print: ISSN 2356–9751**
Vol. (7) Issue (12) (2022), (63-75) **online: ISSN 2356–976x**
http://bjas.journals.ekb.eg

Molecular diagnosis of Type I diabetes in some Egyptian patients using HLA-DRB1 exon 2 gene

Kholud Z.Higazy[1], Amal M.Abdel-kareim[1], Naglaa F.Al-Husseini[2] and Mohammed H.Awwad[1]
[1]Zoology Dept., Faculty of Science, Benha University, Egypt
[2]Biochemistry Dept., Faculty of medicine, Benha University, Egypt
E-mail: kholud.hegazy33@gmail.com.

Abstract

Type 1 diabetes is regarded as a chronic autoimmune disease characterized by dysfunctional T-lymphocyte activation that targets pancreatic Langerhans β-cells, causing insulin deficiency. By raising the amounts of sugar in the blood and urine, this condition causes hyperglycemia. Five groups were chosen for the current experiment, with N standing in for healthy individuals. Four groups of patients were created based on their ages. The bidirectional sequence of the HLA-DRB1 exon 2 gene was determined using blood samples from the five groups. PCR was used to define the gene sequence, amplify it, and isolate the gene for DNA sequencing in two directions. To understand more about the isolated gene and its role in T1DM, bioinformatics analysis of the sequencing output data was carried out. In addition, certain software programs were used by the study teams to forecast the translated protein and RNA secondary structure. The findings showed that an alteration in the secondary structure of RNA affected protein translation and release in bodily cells. Additionally, based on SNP analysis, the four groups' targeted genes revealed mutations caused by the replacement or indel of certain nucleotides, which change the anticipated translated protein in patient groups in comparison to control.

Keywords: Type 1 diabetes, HLA-DRB1, exon 2, DNA sequencing, autoimmune disease

Introduction

Type 1 diabetes mellitus (T1DM) is one of the most prevalent and complicated autoimmune disorders that primarily affects children [1]. It is defined by the inability of the body to manufacture insulin due to the loss of pancreatic cells [2]. The incidence rate of T1DM differs between nations, between ethnic groups, and within populations and countries. It was anticipated that the number of diabetes patients will grow by nearly 180 million by 2030, from 371.3 million in 2012 to 551.87 million [3]. Despite significant treatment advancements, the incidence of type one diabetes continues to rise and has been very varied among nations [4].

Variation in T1DM incidence throughout the world is quite evident. As it varies from 0.1/100 000 per year in China to over 400/100 000 per year in Finland among children under the age of 15 [5]. Environmental variables have been implicated for a long time in the pathogenesis of T1DM, both as initiators and intensifiers of pancreatic -cell death [6]. Egypt is the greatest contributor to the projected total number of children T1DM cases in the Middle East, accounting for around a quarter of the region's total diabetic cases. In children under 15 years of age, the incidence of diabetes ranges from 1/100,000 per year in Pakistan to 8/100,000 per year in Egypt [7, 8].

Diabetes, type I, is characterized by inadequate insulin synthesis in the body. As a result, people with T1DM require daily insulin treatment to balance the glucose level in the blood and preserve their lives from potentially fatal prenatal complications. The cause of type IDM is currently unknown and cannot be prevented. Although the causes of type I diabetes remain unknown, changes in environmental risk factors and/or viral infections may influence the onset of DM. T1DM is characterized by polyuria, polydipsia, polyphagia, weight loss, abnormalities in eyesight, and weariness [9].

It is now commonly known that both environmental and genetic factors contribute to the autoimmune onslaught in type 1 diabetes, which is extremely complex and dependent on a number of previously unidentified variables [10]. T1DM leads to severe diabetes consequences such as retinopathy, neuropathy, nephropathy, cardiovascular problems, and ulceration by causing extensive tissue and vascular damage. Long-term diabetes individuals are also susceptible to microvascular problems and macrovascular disorders such as coronary artery, heart, and peripheral vascular diseases [11, 12, and 13]. Epidemiologic patterns indicate that demographic, regional, biological, cultural, and other variables (the involvement of infections, early childhood food, vitamin D exposure, environmental contaminants, increased height velocity, obesity, and insulin resistance) influence the distribution of T1DM [14].

Various studies have estimated that the risk of type 1 diabetes is around fifty percent hereditary and fifty percent environmental. Non-HLA candidate genes include insulin [15], CTLA4 [16], PTPN22 [17], IL2RA [18]. The HLA genomic region may be one of the most complex and fascinating genetic regions in the entire human and vertebrate genome in terms of infection [19]. Class I contains HLA-A, HLA-B, and HLA-C, while class II contains HLA-DR, HLA-DQ, and HLA-DP. Specific HLA alleles that dictate amino acid residues in each gene are more common in persons

with T1D than in 'controls' (susceptibility alleles) or significantly less common (protective alleles) than in controls [20].

Human Leukocyte Antigen (HLA) class II genes contributed to around fifty percent of the hereditary risk for type 1 diabetes [21]. DRB1 and DQB1 alleles independently or in combination were related with changed T1D risk, and it was found that both susceptible and protective DRB1, DQA1, and DQB1 alleles were implicated in T1D pathogenesis [22, 23]. The connection of DRB1*03:01:01 DQB1*02:01 and DRB1*04:01:01 DQB1*03: 02 haplotypes is responsible for Bahraini Arabs' increased vulnerability to T1D. [24]. As the frequency of HLA alleles, haplotypes, and genotypes vary significantly by population and ethnicity, population studies have established that the association between HLA and T1D also varies by geographical location and ethnic background [25, 26]. Both the high frequency of another vulnerable haplotype (DRB1*03:01-DQA1*05:01-DQB1*02:01) and the low frequency of a protective haplotype (DRB1*15:01-DQA1*01:02-DQB1*06:02) contribute to the high prevalence of T1DM among Sardinians [27]. In this work, we sought to accurately diagnose T1DM patients based on the HLA-DRB1 exon2 sequence.

Subject and methods:

In the present investigation, individuals aged 3 to 20 years were diagnosed with type 1 diabetes mellitus (T1DM) using clinical blood glucose testing. They were sorted into four groups by age range, with a fifth group consisting of normal individuals.

Molecular technique:

DNA extraction: total cellular DNA was isolated from patients' blood by using QIAamp DNA Blood Mini (51104 USA and Canada) as the manufacturer's protocol.

Polymerase chain reaction (PCR): used for gene amplification and detection Taq DNA polymerase was purchased from QIAGEN, GmbH, Germany (cat. nos. 201443) Taq master mix was used for HLA-DRB1 exon2 gene amplification according to the manufacture protocol.

PCR purification: The QIAquick Gel Extraction Kit (cat. nos. 28704) was used for DNA purification from the agarose gel according to the manufacture protocol. The QIAquick Gel Extraction and PCR purification kit effectively removes primers, dNTPS, unincorporated labeled nucleotides, enzymes, and salts from PCR and other reaction mixtures. The HLA-DRB1 exon2 gene primer pair used here included the forward primer DRB1-(F-5' CCCCACAGCACGTTTCTTG 3') and the reverse primer DRB1- (R-5' CCGCTGCACTGTGAAGCTCT 3') according to [28].

DNA Sequencing: BigDye™ Terminator v3.1 Cycle Sequencing Kit was applied to the products of the PCR purification procedure. A common platform for fluorescent dideoxy DNA sequencing is the capillary-based automated DNA sequencer. Capillary-based sequencing uses liquid self-coating polymers to separate DNA fragments and has vastly improved the efficiency of DNA sequencing. Data was obtained using the Big Dye® Terminator 3.1 Sequencing Standard. POP-4 performance Optimized Polymers were used as the separation matrices. A 10× Genetic Analysis buffer with EDTA, diluted to 1×, was used as the running buffer for POP-4. 3730 Buffer (10×) with EDTA, diluted to 1×, was used for POP-7 runs. All reagents, standards, software programs, and updates were from Applied Biosystems (Foster City, CA, USA) [29]. Sequence data were collected using ABI PRISM®3100 Data Collection™ v3.0, and Data were analyzed with ABPRISM®3100 DNA Sequencing Analysis Software™ v5.1.

Bioinformatics and Statistical Analysis:

Molecular analysis (Bioinformatics) tools:

T-Coffee is a multiple sequence alignment online tool. It was used to compare all sequences in normal and treated groups to detect mutations which appear as single nucleotide polymorphism (http://tcoffee.crg.cat/apps/tcoffee/do:regular) [30]. The Vienna RNA Website [31] (http://rna.tbi.univie.ac.at/.) was used for the prediction of RNA secondary structure by free energy minimization method.

Results:

PCR and sequencing techniques were utilized to find mutations in the partial sequence of the HLA-DRB1 exon2 gene in the four distinct T1DM patient groups. The DNA ladder seen in the first lane was split into bands of differing lengths. While the PCR products of the HLA-DRB1 gene for the various normal and patient groups appear as a single band in each lane with a length of approximately 217-221 base pairs (bp) in the forward direction and 196 base pairs in the reverse direction, respectively. Normal, the groups 1, 2, 3, and 4 were organised in lanes 1 through 5 (Fig. 1).

The PCR products were purified from agarose gel after the relevant bands were excised and sequenced. The partial HLA-DRB1 exon2 sequences of all groups investigated are depicted in figure (2 and 3). In addition to NCBI (NG 029921.1), the sequences of HLA- DRB1 exon 2 in all groups consisted of 217-221 nucleotides in the forward direction and 192 nucleotides in the reverse direction.

The multiply sequence alignment were built by using the online software T- Coffee (MSA) of the HLA-DRB1 (Exon 2) in both directions (Fig.2, 3).It revealed that all studied groups were differentiated in two clusters. The first cluster was NCBI sequence and the normal HLA-DRB1 exon2 while the second cluster was groups A1, A2, A3 and A4.

The sequences of HLA-DRB1 EXON 2 in all groups were of 217-221 nucleotides in the forward direction and 192 nucleotides in the reverse direction addition to NCBI (NG_029921.1) sequence (220 bp). The normal and NCBI sequences of HLA-DRB1 exon2 gene were identical with 220 bp. A1, A2, A3, A4 groups were different (SNPS) when compared to the NCBI ref. and normal sequence in both forward and reverse direction. First the forward direction was changed as the following A/T, C/T, A/G, C/A, G/A, A/T, G/A, A/T, A/G, G/T, G/A, A/T, G/T, G/C, T/C, G/A, G/C, G/A, C/G, A/C and C/T at position 19, 21, 25, 26, 28, 37, 47, 51, 52, 54, 63, 66, 89, 98, 113, 117, 159, 161, 162, 170 and 176, respectively. Also, there missing nucleotide in position 22, 24, 50, 211 and extra nucleotides at position 196. While in the reverse direction there were replacement as G/T, C/T, A/T, C/T, C/G, T/A, G/A, A/T, G/C, G/C, G/T, G/A, A/C and T/A at position 24, 55, 58, 77, 82, 88, 104, 107, 114, 160, 163, 166, 167 and 196 , respectively. Also, there missing nucleotide at position 171 and extra nucleotides at position 186.

The substitution, extra and missing nucleotides caused changing in the arrangement and converted the protein amino acids which able to be produced from the HLA-DRB1 exon2 gene to different amino acids by adding some amino acids as N, E, H, K, F, Y, Y and V in the protein amino acid sequence and they were missing from the normal group and NCBI ref. as showed in Table (1) and Fig. (4) and found in one or more of the patients group in forward direction of the gene partial sequence. As well as, in the reverse direction there were other missing amino acids in normal group as E and M, group A2 was missing F amino acid and M amino acid was found only in group A4 as showed in table (2) and Fig.(5).

RNA is one strand and has a higher capacity for hydrogen bonds than DNA, so it tends to form complex base-pairing interactions. Stems and loops are terms for regions with base pairing and a single strand, respectively. For stems, it require more energy to break the hydrogen bonds. The RNA secondary structure was assessed by RNAfold web service (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold.cgi). The results revealed the minimum free energy (MFE) calculated for RNA secondary structures predicted from consensus RNA sequence transcribed from HLA-DRB1 exon2 in forward direction in A1, A3 and A4 T1DM patients studied groups were slightly lower than that of the NCBI ref. and normal sequences. While, it was slightly higher in A2 group. Also, the RNA secondary structure stability in these studied groups will be different than that of the NCBI ref. and normal sequences. Consequently the rate of the protein translation for A1, A3 and A4 studied groups will be slightly greater than that of the NCBI ref. and the normal sequences, but it could be nonfunctional. While as the secondary structure of A2 would be more stable than the NCBI ref. and the normal group and the protein production rate of would be much more less than the protein production of NCBI ref. and normal.

On the other hand, in the reverse direction of HLA-DRB1 exon 2 gene sequence the MFE in A4 were higher than normal and NCBI reference sequence and the protein production rate would be less the normal but in the other three diabetic groups the minimum free energy were lower the normal and that would lead to increase the protein production but it may be non-functional due to the change of amino acid sequence.

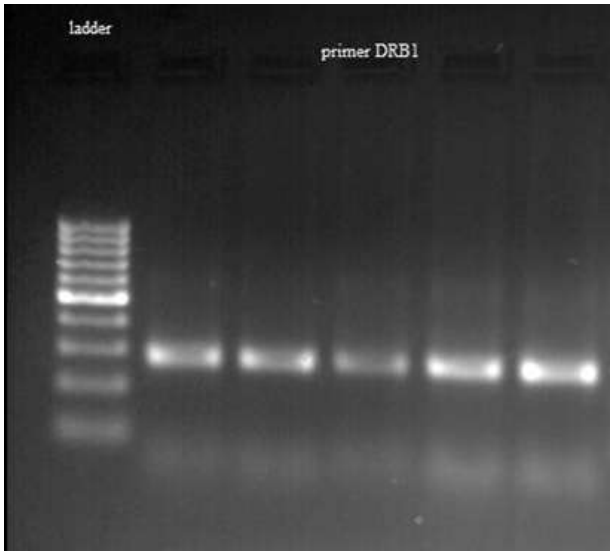

Fig. (1) PCR products of HLA-DRB1 exon2 gene with approximately 220pb

```
NG_029921.1   GGACGGAGCGGGTGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGAACGTGCGCTTCGACAGCG
N             GGACGGAGCGGGTGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGAACGTGCGCTTCGACAGCG
A1            GGACGGAGCGGGTGCGGTACCTGGACAGATACTTCCATAACCAGGAAGA-TGCTTGCGCTTCAACTGCG
A2            GGACGGAGCGGGTGCGGTACCTGGACAGATACTTCCATAACCAGGAAGA-TGCTTGCGCTTCAACTGCG
A3            GGACGGAGCGGGTGCGGTTCT-G-GCAAATACTTCCATAACCAGGAGGAGTACTTGCGCTTCAACTGCG
A4            GGACGGAGCGGGTGCGGTTCCTGGAAAGATACTTCCTTAACCAGGAGGAGAACGTGCGCTTCGACAGCG

cons          ***************** *  *   * ******** ********* **   * ******** ** ***

NG_029921.1   ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGAACAGCCAGAAGG
N             ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGAACAGCCAGAAGG
A1            ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGAACAGCCAGAAGG
A2            ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGAACAGCCAGAAGG
A3            ACGTGGGGGAGTTCCGGGCTGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTGGAACAGCCAGAAGG
A4            ACGTGGGGGAGTTCCGGGCGGTGACGGACCTGGGGCGGCCTGATGCCAAGTACTGGAACAGCCAGAAGG

cons          ****************** ******** ************* *** ********************

NG_029921.1   ACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCAGACACAACTACGGGGT-TGTGGAGAGC
N             ACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCAGACACAACTACGGGGT-TGTGGAGAGC
A1            ACCTCCTGGAGCAGAAGCGGGCCCAGGTGGACACCTACTGCAGACACAACTACGGGGT-TGTGGAGAGC
A2            ACCTCCTGGAGCAGAAGCGGGCCCAGGTGGACACCTACTGCAGACACAACTACGGGGT-TGTGGAGAGC
A3            ACCTCCTGGAGCAGAAGCGGGCCGAGGTGGACACCTACTGCAGACACAACTACGGGGT-TGTGGAGAGC
A4            ACCTCCTGGAGCAGAAGCGGGCCCGGGTGGACCCTTATTGCAGACACAACTACGGGGTTTGTGGAGAGC

cons          ********************* *  *******   ** ******************* **********

NG_029921.1    TTCACAGTGCAGCG
N             TTCACAGTGCAGCG
A1            TTC-CAGTGCAGCG
A2            TTC-CAGTGCAGCG
A3            TTC-CAGTGCAGCG
A4            TTCCCAGTGCAGCG

cons          *** **********
```

Fig. (2) T-coffee alignment of the HLA-DRB1 exon2 DNA consensus sequence of all studied groups in forward direction

```
NG_029921.1     TCCTTCTGGCTGTTCCAGTACTCGGCATCAGGCCGCCCCAGCTCCGTCACCGCCCGGAACTCCCCCACGTCGCTG
N               TCCTTCTGGCTGTTCCAGTACTCGGCATCAGGCCGCCCCAGCTCCGTCACCGCCCGGAACTCCCCCACGTCGCTG
A1              TCCTTCTGGCTGTTCCAGTACTCGGCATCAGGCCGCCCCAGCTCCGTCACCGCCCGGAACTCCCCCACGTCGCTG
A2              TCCTTCTGGCTGTTCCAGTACTCGGCATCAGGCCGCCCCAGCTCCCTCACCGCCCGGAACTCCCCCACGTCGCTG
A3              TCCTTCTGGCTGTTCCAGTACTCGGCATCAGGCCGCCCCAGCTCCGTCACCGCCTGGTACTCCCCCACGTCGCTG
A4              TCCTTCTGGCTGTTCCAGTACTCTGCATCAGGCCGCCCCAGCTCCGTCACCGCCTGGAACTCCCCCACGTCGCTG

cons            ********************** ****************************** ** ******************

NG_029921.1     TCGAAGCGCACGTTCTCCTCCTGGTTATGGAAGTATCTGTCCAGGTACCGCACCCGCTCCGTCCCATTGAAGAAA
N               TCGAAGCGCACGTTCTCCTCCTGGTTATGGAAGTATCTGTCCAGGTACCGCACCCGCTCCGTCCCATTGAAGAAA
A1              TCGAAGCGCACGTACTCCTCTTGGTTATAGAAGTATCTGTCCAGGTACCGCACCCGCTCCGTCCCATTGAAGAAA
A2              TCGAAGCGCACGTTCTCCTCCTGGTTATGGAAGTATCTCTCCAGGTACCGCACCCGCTCCGTCCCATTGAAGAAA
A3              TTGAAGCGCACGTACTCCTCTTGGTTATAGAAGTATCTGTCCAGGTACCGCACCCGCTCCGTCCCATTGAAGAAA
A4              TCGAAGGGCACGTACTCCTCCTGGTTATGGATGTATCTGTCCAGGTACCGCACCCGCTCCGTCCCATTGAAGAAA

cons            * **** ****** ****** ******* ** ****** ***********************************

NG_029921.1     TGACACTCAGACGTAGACTA-CTCCAAGAAACGTGGCTGTGGG
N               TGACACTCAGACGTAGACTA-CTCCAAGAAACGTGGCTGTGGG
A1              TGACACTCAGACTTAACCTAGCTCCAAGAAACGTG-CTGTGGG
A2              TGACACTCAGACTTAAGCAAGCTCCAAGAAACGTG-CTGTGGG
A3              TGACACTCACACTTAACCTAGCTCCAAGAAACGTG-CTGTGGG
A4              TGACACTCACACTTAAACTAGCTCCAAGAAACGTG-CTGTGGG

cons            ********* ** **  * * ************** *******
```

Fig. (3) T-coffee alignment of the HLA-DRB1 exon2 DNA consensus sequence of all studied groups in reverse direction

```
NG_029921.1   GRSGCGTWTDTSITRRRTCASTATWGSSGRRSWGGLMPSTGTARRTSWSRSGAGWTTTADTTTGLWRAS
N             GRSGCGTWTDTSITRRRTCASTATWGSSGRRSWGGLMPSTGTARRTSWSRSGAGWTTTADTTTGLWRAS
A1            GRSGCGTWTDTSITRKMLALQLRRGGVPGGDGAGAACRVLEQPEGPPGAEAGPGGHLLQ-TQLRGCGEL
A2            GRSGCGTWTDTSITRKMLALQLRRGGVPGGDGAGAACRVLEQPEDLLEQKRAQVDTYCR-HNYGVVESF
A3            GRSGCGSGKYFHNQEEYLRFNCDVGEFRAVTELGRPDAEYWNSQRPPGAEAGRGGHLLQ-TQLRGCGEL
A4            GRSGCGSWKDTSLTRRRTCASTATWGSSGRRTWGGLMPSTGTARRTSWSRSGPGWTLIADTTTGFVESF

cons          ******: .     ..    .       .    *        ..     . .

NG_029921.1   QCQ
N             QCQ
A1            PVQ
A2            QCQ
A3            PVQ
A4            PVQ

cons           *
```

Fig. (4) Encoded amino acids of HLA-DRB1 exon2 in forward direction in all groups

```
NG_029921.1         SFWLFQYSASGRPSSVTARNSPTSLSKRTFSSWLWKYLSRYRTRSVPLKKHSDVDYSKKRGCG
N                   SFWLFQYSASGRPSSVTARNSPTSLSKRTFSSWLWKYLSRYRTRSVPLKKHSDVDYSKKRGCG
A1                  SFWLFQYSASGRPSSVTARNSPTSLSKRTYSSWL-KYLSRYRTRSVPLKKHSDLT-LQETCCG
A2                  SFWLFQYSASGRPSSLTARNSPTSLSKRTFSSWLWKYLSRYRTRSVPLKKHSDLSKLQETCCG
A3                  SFWLFQYSASGRPSSVTAWYSPTSLLKRTYSSWL-KYLSRYRTRSVPLKKHSHLT-LQETCCG
A4                  SFWLFQYSASGRPSSVTAWNSPTSLSKGTYSSWLWMYLSRYRTRSVPLKKHSHLN-LQETCCG
cons           ***************:**  ***** * *:****  ****************.:   ::  **
```

Fig. (5) Encoded amino acids of HLA-DRB1 exon2 in reverse direction in all groups

Table (1) percentage and numbers of amino acids in protein of HLA-DRB1 exon2 in forward direction for all groups

Amino Acids (F)	Normal %	A1 %	A2 %	A3 %	A4 %
Ala (A)	6 8.33%	6 8.33	5 6.94	4 5.56	4 5.56
Arg (R)	10 13.89%	6 8.33	7 9.72	7 9.72	9 12.50
Asn (N)	--------	---	1 1.39	3 4.17	----
Asp (D)	2 2.78%	2 2.78	4 5.56	2 2.78	2 2.78
Cys (C)	3 4.17%	3 4.17	4 5.56	3 4.17	2 2.78
Gln (Q)	2 2.78%	5 6.94	6 8.33	5 6.94	1 1.39
Glu (E)	-------	4 5.56	4 5.56	7 9.72	1 1.39
Gly (G)	11 15.28	16 22.22	10 13.89	12 16.67	11 15.28
His (H)	---	1 1.39	1 1.39	2 2.78	---
Ile (I)	1 1.39	1 1.39	1 1.39	----	1 1.39
Leu (L)	2 2.78	8 11.11	6 8.33	6 8.33	3 4.17
Lys (k)	---	1 1.39	2 2.78	1 1.39	1 1.39
Met (M)	1 1.39	1 1.39	1 1.39	----	1 1.39
Phe (F)	----	----	1 1.39	3 4.17	2 2.78
Pro (P)	1 1.39	6 8.33	2 2.78	4 5.56	3 4.17
Ser (S)	11 15.28	2 2.78	3 4.17	3 4.17	11 15.28
Thr (T)	16 22.22	5 6.94	5 6.94	2 2.78	13 18.06
Trp (W)	6 8.33	1 1.39	1 1.39	1 1.39	5 6.94
Tyr (Y)	----	---	2 2.78	3 4.17	----
Val (V)	---	3 4.17	5 6.94	3 4.17	2 2.78

Table (2) percentage and numbers of amino acids in protein of HLA-DRB1 exon2 in reverse direction for all groups.

Amino acids (R)	Normal %	A1 %	A2 %	A3 %	A4 %
Ala	2	2	2	2	2
(A)	3.17	3.17	3.17	3.17	3.17
Arg	7	6	6	5	4
(R)	11.11	9.52	9.52	7.94	6.35
Asn	1	1	1	---	2
(N)	1.59	1.59	1.59		3.17
Asp	2	1	1	---	---
(D)	3.17	1.59	1.59		
Cys	1	2	2	2	2
(C)	1.59	3.17	3.17	3.17	3.17
Gln	1	2	2	2	2
(Q)	1.59	3.17	3.17	3.17	3.17
Glu	---	1	1	1	1
(E)		1.59	1.59	1.59	1.59
Gly	3	2	2	2	3
(G)	4.76	3.17	3.17	3.17	4.76
His	1	1	1	2	2
(H)	1.59	1.59	1.59	3.17	3.17
Ile	---	---	---	---	----
(I)					
Leu	5	7	8	8	7
(L)	7.94	11.11	12.70	12.70	11.11
Lys	6	4	5	4	3
(k)	9.52	6.35	7.94	6.35	4.76
Met	---	---	---	---	1
(M)					1.59
Phe	3	2	---	2	2
(F)	4.76	3.17		3.17	3.17
Pro	3	3	3	3	3
(P)	4.76	4.76	4.76	4.76	4.76
Ser	14	13	14	12	13
(S)	22.22	20.63	22.22	19.05	20.63
Thr	4	6	5	6	5
(T)	6.35	9.52	7.94	9.52	7.94
Trp	3	2	3	3	4
(W)	4.76	3.17	4.76	4.76	6.35
Tyr	4	4	3	5	4
(Y)	6.35	6.35	4.76	7.94	6.35
Val	3	2	1	2	2
(V)	4.76	3.17	1.59	3.17	3.17

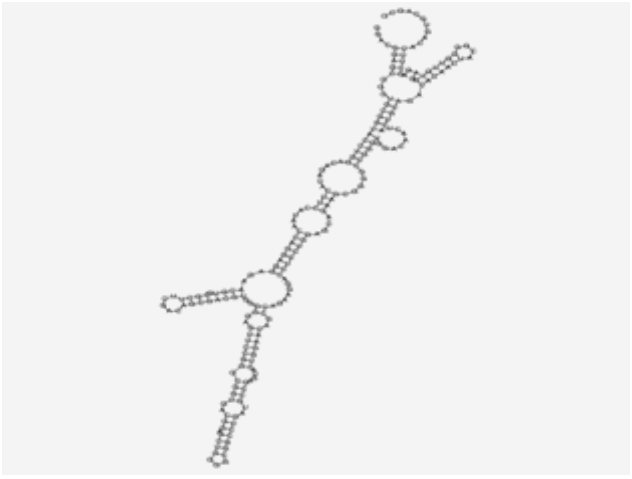

Fig. (6) The predicted RNA secondary structure for normal group and NCBI (NG_029921.1) Ref. sequence in forward direction. Free energy-75.70 kcal/mol

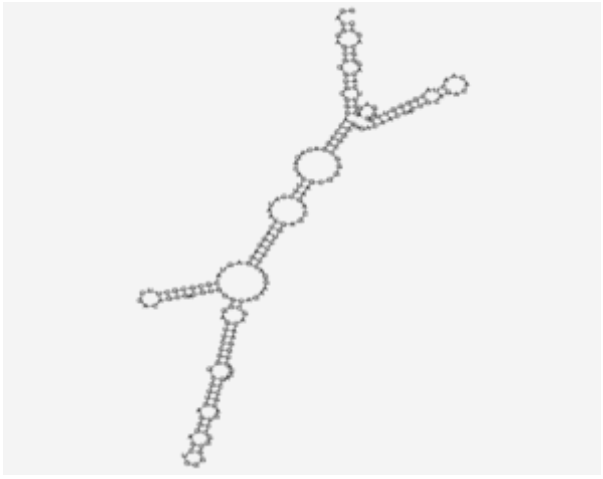

Fig. (7) The predicted RNA secondary structure for A1 group sequence in forward direction. Free energy -81.60 kcal/mol

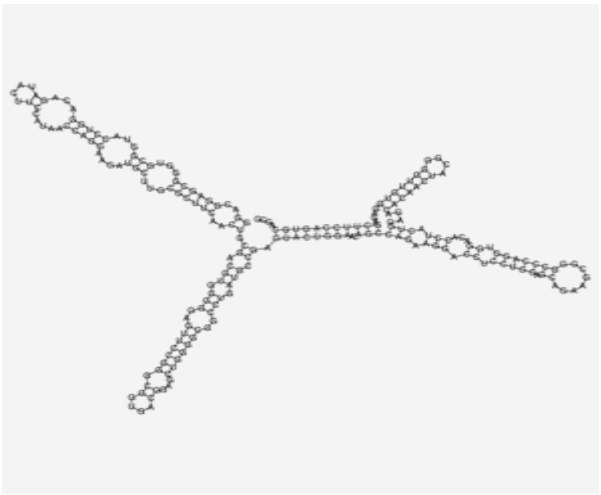

Fig. (8) The predicted RNA secondary structure for A2 group sequence in forward direction. Free energy -72.50 kcal/mol

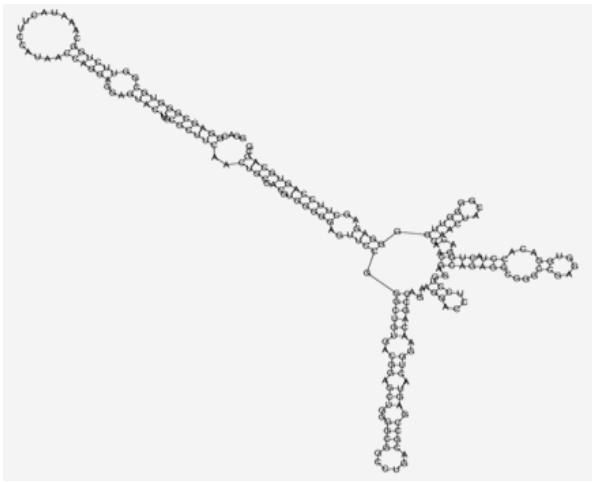

Fig. (9) The predicted RNA secondary structure for A3 group sequence in forward direction. Free energy -81.50 kcal/mol

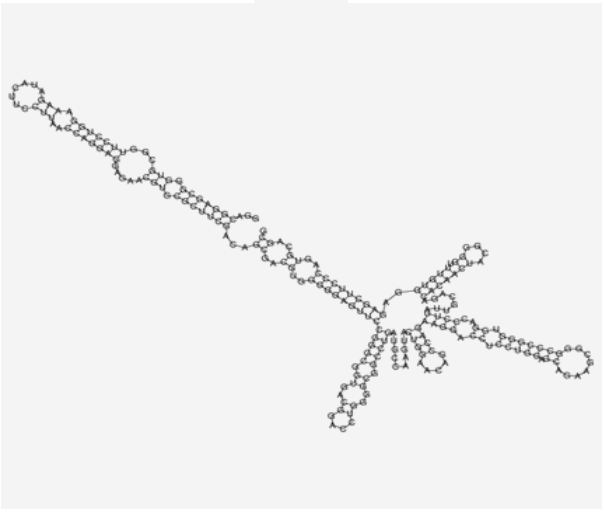

Fig. (10) The predicted RNA secondary structure for A4 group sequence in forward direction. Free energy -77.60 kcal/mol.

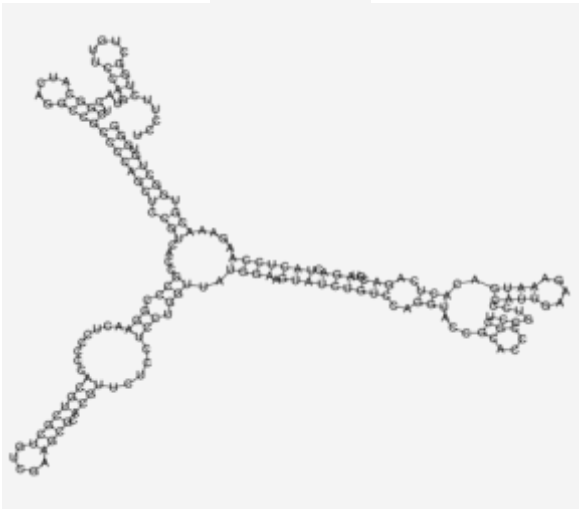

Fig. (11) The predicted RNA secondary structure for normal group and NCBI (NG_029921.1) Ref. sequence in reverse direction. Free energy-46.70 kcal/mol.

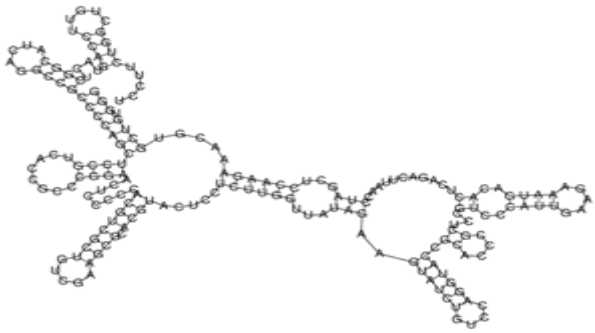

Fig. (12) The predicted RNA secondary structure for A1 group sequence in reverse direction. Free energy -47.20 kcal/mol

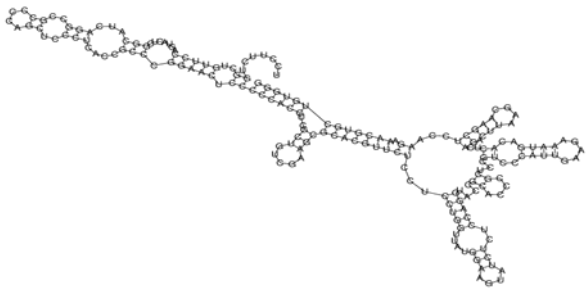

Fig. (13) The predicted RNA secondary structure for A2 group sequence in reverse direction. Free energy -47.23 kcal/mol.

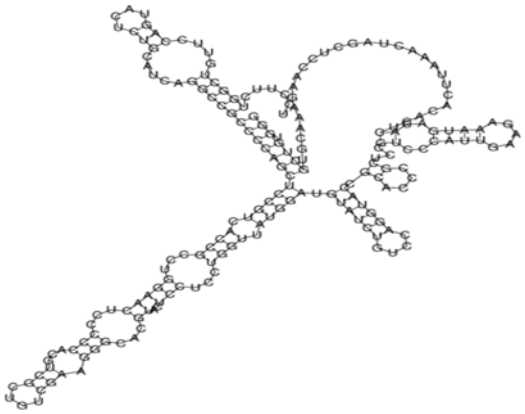

Fig. (14) The predicted RNA secondary structure for A3 group sequence in reverse direction. Free energy -46.90 kcal/mol.

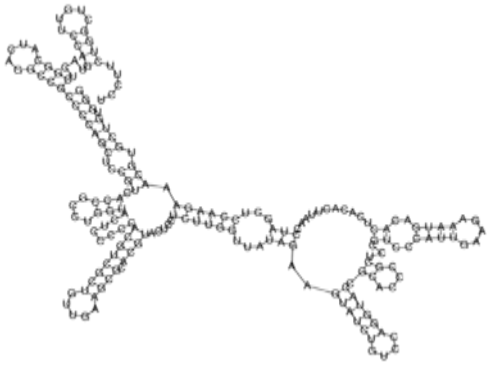

Fig. (15) The predicted RNA secondary structure for A4 group sequence in reverse direction. Free energy -41.70 kcal/mol

Discussion

Several prior genetics studies have found over sixty genes related with the likelihood of acquiring T1DM. Genome-wide association studies (GWAS) identified the majority of genetic loci using large case-control populations of European origin [32]. Understanding the biology of type 1 diabetes has been fundamentally aided by genetic investigations [33]. Human leukocyte antigen (HLA) regions were the first to be associated with Type 1 diabetes [34, 35, 36] As a result of this finding, scientists have attempted to comprehend the underlying processes through which alleles of HLA-encoding genes contribute to the T1D association.

Recent molecular research revealed discrepancies in the partial HLA-DRB1 exon2 gene sequence between T1DM patients and NCBI database sequences. In addition, they identify the mutations that differentiate T1DM groups from non-diabetic ones. In this investigation, total DNA was taken from the patient's blood tissue, and then the targeted sequences were amplified using the Polymerase Chain Reaction (PCR) technique. The target bands at roughly 221 bp in the forward direction and 196 bp in the reverse direction were visualised by electrophoresis, excised from the gel, and purified. DNA sequencing and bioinformatics tools such as T-coffee, RNAfold web server, and other bioinformatics tools were utilised to evaluate genetic variants in the targeted regions of the HLA-DRB1 gene.

The Sanger DNA sequencing technique is one of the fundamental methods used to identify the precise sequence of nucleotides inside a DNA molecule. This technique is able to analyze thousands of genes by high throughput sequencing, and there are a number of issues that currently urge the implementation of high throughput sequencing in routine clinical diagnostics due to its accuracy, genotype-phenotype correlation, clinical utility, and ethical considerations [37]. In this work, this approach was utilized to explore the partial sequence of the HLA-DRB1 exon 2 gene in a bidirectional manner for mutations that occurred in T1DM patients relative to the normal population and NCBI reference. Without bioinformatics tools,

the sequencing of raw materials would not produce high-quality findings; hence, these tools are required to get the desired outcome [38].

Koonin stated that the biological sciences are composed of three groups. DNA stores genetic information, RNA translates to proteins, and proteins are important for several biological processes [39]. Therefore, it was suitable to utilize Sequence Manipulation Suite to predict the RNA and the translated protein sequence for further research.

SNPs are the most prevalent kind of genetic variation and serve as the foundation for the majority of molecular markers. Before they can be employed for direct sequence-based SNP detection or in a derived SNP assay, however, they must be recognised. Multiple sequence fragments originating from distinct genotypes and reflecting the same area of the genome can be aligned in order to identify sequence variations. The corresponding mismatched nucleotides are likely SNPs or insertions/deletions [40]. Exon2 of the HLA-DRB1 gene in the A1, A2, A3, and A4 patients included a significant number of missing, extra-addition, and replacement nucleotides. These mutations alter the predicted amino acid sequences of their patients' HLA-DRB1 exon2 proteins. RNA molecules, which are naturally single-stranded, fold onto themselves to form physiologically active 3D structures, using Watson-Crick base pairing principles that are mostly comparable [41]. Which has a regular pattern of Watson-Crick and GU pairs (helices) and less regularly arranged lengths of nucleotides (loops) [42].

RNA secondary structure and minimal free energy (MFE) evaluation may be performed using a variety of techniques and methods for structure prediction [43]. RNA secondary structure plays an important role in the biosynthesis of proteins, and it can have a negative effect on translation by slowing or blocking the initiation and movement of ribosomes along the mRNA, thereby reducing the yield of protein and becoming a major factor in the regulation of gene expression by altering the minimum free energy required to unwind the secondary structure stems and initiate the translation process [44]. In this work, the RNA secondary structure was reconstructed using the RNAFOLD web server. The results showed that the computed free energy for RNA secondary structures predicted from consensus RNA sequence transcribed from HLA-DRB1 exon2 in forward direction was somewhat less than that of the NCBI reference and normal sequences. However, it was slightly greater in the A2 group. In addition, the stability of the secondary structure of RNA in these examined groups will differ from that of the NCBI reference and normal sequences. Therefore, the protein translation rate for the A1, A3, and A4 examined groups will be slightly higher than that of the NCBI reference and the usual sequences, but it may be dysfunctional. While A2's secondary structure would be more stable than that of the NCBI reference and the normal group, its protein synthesis rate would be significantly lower than that of the NCBI reference and the normal group.

On the other hand, in the reverse direction of the HLA-DRB1 exon 2 gene sequence, the MFE in A4 was greater than normal and the NCBI reference sequence, and the protein production rate would be lower than normal. However, in the other three diabetic groups, the minimum free energy was lower than normal, which would lead to an increase in protein production, although it may be non-functional due to a change in amino acid sequence.

In conclusion, this study showed that the HLA-DRB1 exon2 gene of the four studied groups were mutated when compared to the normal and NCBI reference.

Reference

[1] Pang, H., Luo, S., Xiao, Y., Xia, Y., Li, X., Huang, G. & Zhou, Z. (2020). Emerging roles of exosomes in T1DM. Frontiers in Immunology, 11, 593348.

[2] Aathira, R., & Jain, V. (2014). Advances in management of type 1 diabetes mellitus. World journal of diabetes, 5(5), 689.

[3] Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care. 2004 May; 27(5):1047-53.

[4] Jahromi, M., & Al-Ozairi, E. (2019). Human leukocyte antigen (HLA) and islet autoantibodies are tools to characterize type 1 diabetes in Arab countries: emphasis on Kuwait. Disease Markers, 2019.

[5] Eurodiab Ace Study Group. (2000). Variation and trends in incidence of childhood diabetes in Europe. The Lancet, 355(9207), 873-876.

[6] Knip, M., & Simell, O. (2012). Environmental triggers of type 1 diabetes. Cold Spring Harbor perspectives in medicine, 2(7), a007690.

[7] Atlas, D. (2015). International diabetes federation. IDF Diabetes Atlas, 7th edn. Brussels, Belgium: International Diabetes Federation, 33.

[8] El-Ziny, M. A. E. M., Salem, N. A. B., El-Hawary, A. K., Chalaby, N. M., & Elsharkawy, A. A. E. (2014). Epidemiology of childhood type 1 diabetes mellitus in Nile Delta, northern Egypt-a retrospective study. Journal of clinical research in pediatric endocrinology, 6(1), 9.

[9] Okur, M. E., Karantas, I. D., & Siafaka, P. I. (2017). Diabetes mellitus: A review on

pathophysiology, current status of oral medications and future perspectives.

[10] Dunger, D. B., Sperling, M. A., Acerini, C. L., Bohn, D. J., Daneman, D., Danne, T., ... & Wolfsdorf, J. I. (2004). European Society for Paediatric Endocrinology/Lawson Wilkins Pediatric Endocrine Society consensus statement on diabetic ketoacidosis in children and adolescents. Pediatrics, 113(2), e133

[11] Svensson, M., Eriksson, J. W., & Dahlquist, G. (2004). Early glycemic control, age at onset, and development of microvascular complications in childhood-onset type 1 diabetes: a population-based study in northern Sweden. Diabetes care, 27(4), 955-962..

[12] Saely, C. H., Aczel, S., Marte, T., Langer, P., & Drexel, H. (2004). Cardiovascular complications in Type 2 diabetes mellitus depend on the coronary angiographic state rather than on the diabetic state. Diabetologia, 47(1), 145-146.

[13] Deshmukh, C. D., Jain, A., & Nahata, B. (2015). Diabetes mellitus: a review. Int. J. Pure Appl. Biosci, 3(3), 224-230.

[14] Maahs, D. M., West, N. A., Lawrence, J. M., & Mayer-Davis, E. J. (2010). Epidemiology of type 1 diabetes. Endocrinology and Metabolism Clinics, 39(3), 481-497.

[15] Barratt, B. J., Payne, F., Lowe, C. E., Hermann, R., Healy, B. C., Harold, D. ... & Todd, J. A. (2004). Remapping the insulin gene/IDDM2 locus in type 1 diabetes. Diabetes, 53(7), 1884-1889.

[16] Nisticò, L., Buzzetti, R., Pritchard, L. E., Van der Auwera, B., Giovannini, C., Bosi, E., ... & Todd, J. A. (1996). The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Human molecular genetics, 5(7), 1075-1080.

[17] Bottini, N., Musumeci, L., Alonso, A., Rahmouni, S., Nika, K., Rostamkhani, M. ... & Mustelin, T. (2004). A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nature genetics, 36(4), 337-338.

[18] Vella, A., Cooper, J. D., Lowe, C. E., Walker, N., Nutland, S., Widmer, B. ... & Todd, J. A. (2005). Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. The American Journal of Human Genetics, 76(5), 773-779.

[19] Horton, R., Wilming, L., Rand, V., Lovering, R. C., Bruford, E. A., Khodiyar, V. K., ... & Beck, S. (2004). Gene map of the extended human MHC. Nature Reviews Genetics, 5(12), 889-899.

[20] Rich, S. S. (2017). Genetics and its potential to improve T1D care. Current opinion in endocrinology, diabetes, and obesity, 24(4), 279.

[21] Pociot, F., Akolkar, B., Concannon, P., Erlich, H. A., Julier, C., Morahan, G. ... & Nerup, J. (2010). Genetics of type 1 diabetes: what's next? Diabetes, 59(7), 1561-1571.

[22] Noble, J. A., Valdes, A. M., Cook, M., Klitz, W., Thomson, G., & Erlich, H. A. (1996). The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. American journal of human genetics, 59(5), 1134.

[23] Erlich, H., Valdes, A. M., Noble, J., Carlson, J. A., Varney, M., Concannon, P. ... & Type 1 Diabetes Genetics Consortium. (2008). HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. Diabetes, 57(4), 1084-1092.

[24] Al-Jenaidi, F. A., Wakim-Ghorayeb, S. F., Al-Abbasi, A., Arekat, M. R., Irani-Hakime, N., Najm, P. ... & Almawi, W. Y. (2005). Contribution of selective HLA-DRB1/DQB1 alleles and haplotypes to the genetic susceptibility of type 1 diabetes among Lebanese and Bahraini Arabs. The Journal of Clinical Endocrinology & Metabolism, 90(9), 5104-5109.

[25] Zabeen, B., Govender, D., Hassan, Z., Noble, J. A., Lane, J. A., Mack, S. J., ... & Ogle, G. D. (2019). Clinical features, biochemistry and HLA-DRB1 status in children and adolescents with diabetes in Dhaka, Bangladesh. Diabetes research and clinical practice, 158, 107894.

[26] Al Yafei, Z., Mack, S. J., Alvares, M., Ali, B. R., Afandi, B., Beshyah, S. A., ... & ElGhazali, G. (2022). HLA-DRB1 and–DQB1 Alleles, Haplotypes and Genotypes in Emirati Patients with Type 1 Diabetes Underscores the Benefits of Evaluating Understudied Populations. Frontiers in genetics, 13.

[27] Nejentsev, S., Koskinen, S., Sjöroos, M., Reijonen, H., Schwartz, E. I., Kovalchuk, L., ... & Ilonen, J. (1998). Distribution of insulin-dependent diabetes mellitus (IDDM) - related HLA alleles correlates with the difference in IDDM incidence in four populations of the Eastern Baltic region. Tissue Antigens, 52(5), 473-477.

[28] Raha, O., Sarkar, B., Lakkakula, B. V., Pasumarthy, V., Godi, S., Chowdhury, S., & Vadlamudi, R. R. (2013). HLA class II SNP interactions and the association with type 1 diabetes mellitus in Bengali speaking

patients of Eastern India. Journal of Biomedical Science, 20(1), 1-6.
[29] Detwiler M, Hamp T and Kazim L (2004): "DNA sequencing using the liquid polymer POP-7™ on an ABI PRISM® 3100 Genetic Analyzer." Biotechniques. 36(6):932-933.
[30] Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of molecular biology, 302(1), 205-217.
[31] Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., & Hofacker, I. L. (2008). The vienna RNA websuite. Nucleic acids research, 36(suppl_2), W70-W74.
[32] Pociot, F. (2017). Type 1 diabetes genome-wide association studies: not to be lost in translation. Clinical & translational immunology, 6(12), e162.
[33] Polychronakos, C., & Li, Q. (2011). Understanding type 1 diabetes through genetics: advances and prospects. Nature Reviews Genetics, 12(11), 781-79
[34] Cudworth, A. G., & Woodrow, J. C. (1974). HLA antigens and diabetes mellitus. The Lancet, 304(7889), 1153.
[35] Nerup, J., Platz, P., Andersen, O. O., Christy, M., Lyngsøe, J., Poulsen, J. E., ... & Svejgaard, A. (1974). HLA antigens and diabetes mellitus. The Lancet, 304(7885), 864-866.
[36] Singal, D. P., & Blajchman, M. A. (1973). Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. Diabetes, 22(6), 429-432.
[37] Tipu, H. N., & Shabbir, A. (2015). Evolution of DNA sequencing. J Coll Physicians Surg Pak, 25(3), 210-5.
[38] Mardis, E. R. (2011). A decade's perspective on DNA sequencing technology. Nature, 470(7333), 198-203.
[39] Koonin, E.V. Does the central dogma still stand?. Biol Direct 7, 27 (2012). https://doi.org/10.1186/1745-6150-7-27
[40] Oeveren, J. V., & Janssen, A. (2009). Mining SNPs from DNA sequence data; computational approaches to SNP discovery and analysis. In Single Nucleotide Polymorphisms (pp. 73-91). Humana Press, Totowa, NJ.
[41] Saman Booy, M., Ilin, A., & Orponen, P. (2022). RNA secondary structure prediction with convolutional neural networks. BMC bioinformatics, 23(1), 1-15.
[42] Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA Package 2.0. Algorithms for molecular biology, 6(1), 1-14.
[43] Knudsen, B., & Hein, J. (2003). Pfold: RNA secondary structure prediction using stochastic context-free grammars. Nucleic acids research, 31 (13), 3423-3428.
[44] Gaspar, P., Moura, G., Santos, M. A., & Oliveira, J. L. (2013). mRNA secondary structure optimization using a correlated stem–loop prediction Nucleic acids research, 41 (6), e73-e73.