Journal of Basic and Environmental Sciences

ISSN
Online: 2356-6388
Print: 2536-9202

Research paper open access

Genetic investigations on certain genes associated with type 1 diabetes in some Egyptions

Noha A.El kamshishy, Safaa Alzawahry and Mohammed H.Awwad

Zoology Dept., Faculty of Science, Benha University, Egypt

E-mail : nohaelkamshishy@gmail.com

Abstract

Type 1 diabetes mellitus (T1DM) is an autoimmune disorder that typically manifests in early childhood but can occur in persons of any age. The condition is a result of autoreactive T lymphocytes that target and destroy insulin-producing beta cells in the pancreatic islet. This results in a complete lack of insulin and persistently high blood sugar levels, necessitating lifelong insulin treatment. For the present experiment, five groups were selected, all of whom were volunteers. Four cohorts of T1DM patients were categorized based on their age, whereas the fifth cohort consisted of individuals without diabetes. The HLA-DRB1*03 exon 2 allele's bidirectional sequencing was established using blood samples. PCR was employed to determine the gene sequence, amplify it, and extract the gene for DNA sequencing in both forward and reverse directions. The sequencing output data was subjected to bioinformatics analysis to get insights into the isolated gene and its function in T1DM. In addition, specific software applications were utilized to forecast the translated protein and RNA secondary structures. A modification occurred in the secondary structure of RNA, which had an impact on the translation and release of proteins in somatic cells. Furthermore, the SNP analysis indicated that the targeted genes in the four groups exhibited changes resulting from the substitution or deletion of certain nucleotides. These mutations altered the predicted protein translation in the sick groups as compared to the control group.

1. Introduction

Type 1 diabetes mellitus (T1DM) is a medical illness that occurs when the body's immune system mistakenly attacks and damages the insulin-producing cells in the pancreas. This damage typically results in a significant lack of insulin production inside the body. T1DM comprises around 5-10% of the total number of diabetes cases [1,2] (Quattrin et al., 2023; Holt et al., 2021). T1DM can develop at any stage of life, and the frequency of occurrence differs depending on age [3] (Stene & Tuomilehto, 2024). While the typical age for the appearance of symptoms is usually during childhood or adolescence, there are cases when symptoms might manifest at a later stage [4] (Katsarou et al., 2017). While the highest occurrence of T1DM happens during puberty and early adulthood, it can develop at any age. People with T1DM can live for many years after being diagnosed, which means that the overall number of adults with T1DM is higher than children. This is why our attention is directed towards T1DM in adults [5] (Miller et al., 2012).

The prevalence of T1DM is on the rise globally, with an annual increase of roughly 2–3 percent. This results in tens of thousands of new cases each year worldwide [6,7] (Esposito et al., 2019; Centers for Disease Control and Prevention, 2017). The worldwide occurrence of T1DM is 5.9 cases per 10,000 individuals, with the number of new cases increasing significantly in the past 50 years. Currently, it is anticipated that there are 15 new cases per 100,000 individuals per year [2,8] (Holt et al., 2021; Mobasseri et al., 2020).

Furthermore, the International Diabetes Federation (IDF) estimates that there are 1.1 million individuals between the ages of 14 and 19 who have T1DM. If no measures are taken to stop the rise in diabetes, the number of individuals affected by it will reach a minimum of 629 million by the year 2045. High levels of blood glucose result in about 4 million fatalities annually, and according to the International Diabetes Federation (2017), the estimated worldwide expenditure on diabetes healthcare for adults in 2017 was US$850 billion [9].

Global studies have demonstrated that the prevalence of T1DM has been steadily rising at a rate of 2-5%. Furthermore, the incidence of T1DM in the United States is roughly 1 in 300 among individuals who are 18 years old. In recent times, there has been a shift in the perception that T1DM mostly affects children and teenagers. The age at which symptoms appear is no longer seen as a significant issue as it was previously

believed [10] (Shojaeian and Mehri-Ghahfarrokhi, 2018).

Among the global rankings for diabetes prevalence, six nations in the Middle East and North Africa (MENA) area are included in the top 10. The nations mentioned are Bahrain, Egypt, Kuwait, Oman, Saudi Arabia, and the United Arab Emirates [11] (International Diabetes Federation, 2015). Egypt makes the highest contribution to the overall number of estimated childhoods T1DM cases among Eastern Mediterranean and Middle Eastern countries, accounting for around 25% of the region's total [12] (International Diabetes Federation, 2012). The prevalence ranges from 1/100000 annually in Pakistan to 8/100000 annually in Egypt among children under 15 years of age [13] (Soltèsz et al., 2006).

In their 2014 study, El-Ziny et al. conducted the initial pediatric T1DM epidemiology investigation in the Nile Delta region of Egypt and found a rise in the occurrence and frequency of T1DM over a span of 18 years (1994–2011). There was a higher occurrence and frequency of cases in females, and a greater number of cases were identified as originating from rural areas[14] (El-Ziny ., 2014). There is a lack of epidemiological studies on childhood T1DM in Egypt. Several factors have been identified as contributing to this phenomenon, such as the absence of diabetes registries, the presence of dispersed medical facilities, and inadequate identification of new cases [15] (Attia et al., 2023).

The majority (95%) of cases of T1DM are caused by environmental variables that interact with an individual who is genetically predisposed. Previous research on T1DM has found evidence of genetic and environmental risk factors, as indicated by early familial aggregation and twins' investigations [16] (Redono & Fain, 2001). The presence of a genetic predisposition to T1D, although not fully understood, is evident. Identical twins show a concordance rate of 60%–70% with long-term monitoring [17] (Redondo et al., 2008), whereas first-degree relatives have a lifetime risk of T1D of 5%–6%.

Genetic analyses have shown that the HLA complex, particularly class II alleles, is associated with an increased risk of developing T1DM. This link has been established in studies conducted by Jones et al. (2006), Noble et al. (1996), Morel et al. (1988), and Todd et al. (1987) [18-21]. Class II HLA alleles (DR, DQ, and others) are responsible for about 50% of the hereditary risk, as stated by [18,19] Jones et al. in 2006 and Noble et al. in 1996. Various ethnic groups have identified possible genetic risk loci for T1DM, such

as HLA class I molecules [22] (Noble et al., 2010), as well as non-HLA genes including CTLA4, PTPN22, INS, and IL2RA [23-25] (Besser et al., 2022; Bonifacio et al., 2018; Polychronakos & Li., 2011).

Extensive research over a period of forty years has demonstrated that immunological genes, particularly those responsible for encoding classical human leukocyte antigens (HLA), are associated with the highest genetic susceptibility to disease [26,27] (Noble et al., 2024; Noble, 2015). The correlation between human leukocyte antigens (HLA) and type 1 diabetes was initially documented in the 1970s [28-30] (Cudworth and Woodrow, 1974; Nerup et al., 1974; Singal and Blajchman, 1973). Explaining the HLA genes might be challenging due to their complexity and intricacy. HLA is divided into two classes. The Class I HLA molecules, specifically HLA-A, HLA-B, and HLA-C, are each encoded by a single gene. Class II antigens are formed by a heterodimer consisting of two genes. The HLA-DQA1 gene encodes the chain of the DQ antigen, the HLA-DPA1 gene encodes the chain of the DP antigen, and the HLA-DRA1 gene encodes the chain of the DR antigen [27] (Noble, 2015).

Approximately 25 years ago, researchers noticed that the frequencies of specific HLA antigens associated with susceptibility to T1DM were higher, while the frequencies of protective antigens were lower, in patients with T1DM compared to unrelated healthy individuals of the same ethnicity[31,29] (Rich., 2017; Nerup et al.,1974).The susceptibility to and protection against the development of T1DM are primarily linked to the highly variable exon 2 sequences of the HLA class II genes: DQB1, DQA1, and DRB1[32-39] (Noble and Valdes, 2011; Varney et al., 2010; Howson et al., 2009; Erlich et al., 2008; Nejentsev et al., 2007; Johansson et al., 2003; Thorsby, 1997; She., 1996).

HLA genes have the highest level of genetic variation among all genes in the human genome. While the majority of genes are either monomorphic or contain only a few variant sequences, classical HLA genes have the potential to possess thousands of variable sequences [40,41] (Robinson et al., 2015; González-Galarza et al., 2014). Due to the recognition of the substantial degree of HLA variation and advancements in molecular techniques, multiple molecular HLA typing methods have been created [42-44] (Moyer and Gandhi, 2022; Bravo-Egana et al., 2021; Dunn, 2011). The capacity of these class II molecules to exhibit antigens relies, to some extent, on the amino acid

composition of their alpha and beta chains. Modifications at one or two crucial locations can have a substantial impact on the interaction between key autoantigens, leading to an increase or decrease in the likelihood of developing T1DM [45] (Rowe et al., 1994). HLA alleles exhibit significant variation in frequency and combinations among many ethnic populations [46] (Dashti et al., 2023). The objective of this study was to precisely diagnose individuals with T1DM by analyzing the sequence of exon 2 of the HLA-DRB1*03 allele.

2. Subject and methods:

This experiment involved the participation of volunteers aged 3 to 20 years who were diagnosed with T1DM utilizing rapid, 2-hour postprandial blood glucose testing, and glycosylated hemoglobin levels. The individuals were categorized into four cohorts based on their age, while a fifth cohort was comprised of individuals who were considered to be normal volunteers.

Molecular technique:

DNA extraction: The whole cellular DNA was obtained from the blood of patients using the QIAamp DNA Blood Mini kit (51104 USA and Canada) following the instructions provided by the manufacturer.

The polymerase chain reaction (PCR): is a technique utilized to amplify and detect genes. The Taq DNA polymerase was acquired from QIAGEN, GmbH, Germany (catalog numbers 201443). The amplification of the HLA-DRB1*03 allele exon2 gene was performed using Taq master mix, following the instructions provided by the manufacturer.

PCR purification: was performed using the QIAquick Gel Extraction Kit (cat. nos. 28704) following the manufacturer's procedure to purify DNA from the agarose gel. The QIAquick Gel Extraction and PCR Purification Kit effectively removes primers, dNTPs, unlabeled nucleotides, enzymes, and salts from PCR and other reaction mixtures. The primer pair employed in this study for the HLA-DRB1*03 allele exon2 gene consisted of the forward primer DRB1*03 (F-5’ TACTTCCATAACCAGGAGGAGA 3’) and the reverse primer DRB1*03 (R-5’ TGCAGTAGTTGTCCACCCG 3'), as described by Wu et al. (2013) [47].

DNA Sequencing: The products of the PCR purification technique were subjected to the use of the BigDyeTM Terminator v3.1 Cycle Sequencing Kit. The capillary-based automated DNA sequencer is a widely used platform for

fluorescence dideoxy DNA sequencing. Capillary-based sequencing employs liquid self-coating polymers to separate DNA fragments, resulting in significantly enhanced efficiency in DNA sequencing. The Big Dye® Terminator 3.1 Sequencing Standard was utilized to acquire the data. Performance of the POP-4 Efficiently optimized Polymers served as the matrix for separation. The running buffer for POP-4. 3730 was prepared by diluting a 10× genetic analysis buffer containing EDTA to a 1× concentration. A solution containing EDTA was prepared at a concentration of 10 times the desired final concentration. This solution was then diluted to the desired final concentration and utilized for the POP-7 runs. The reagents, standards, software programs, and updates used in this study were obtained from Applied Biosystems, located in Foster City, CA, USA [48] (Detwiler et al., 2004). The collection of sequence data was performed using the ABI PRISM®3100 Data CollectionTM v3.0, and subsequent analysis of the data was conducted using the ABPRISM®3100 DNA Sequencing Analysis SoftwareTM v5.1.

Bioinformatics and Statistical Analysis:

Molecular analysis (Bioinformatics) tools: The T-Coffee online tool was utilized to perform a multiple sequence alignment and compare the sequences in the normal and treated groups. The purpose was to identify any mutations that manifest as single nucleotide polymorphisms (SNPs) [49] (Noudame et al., 2000). The tool may be accessed at http://tcoffee.crg.cat/apps/tcoffee/do:regul ar. The Vienna RNA Website, developed by Gruber et al. in 2008[50], was utilized to predict RNA secondary structure using the free energy minimization technique. The website may be accessed at http://rna.tbi.univie.ac.at/.

3. 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 270 base pairs (bp) in the forward direction and 270 base pairs in the reverse direction, respectively, Normal, the groups 1, 2, 3, and 4 were organized in lanes 1 through 5 (Fig. 1).

The PCR products were purified from an agarose gel after the relevant bands were excised and sequenced. The partial HLA-DRB1*03 exon2 allele sequences of all groups investigated are

depicted in figures 2 and 3. In addition to NCBI (HE963767.1), the sequences of the HLA-DRB1*03 exon 2 allele in all groups consisted of 270 nucleotides in the forward direction and 270 nucleotides in the reverse direction.

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

In all groups, the HLA-DRB1 EXON 2 sequences were 270 nucleotides in the forward direction and 270 nucleotides in the reverse direction, in addition to the NCBI (NG_029921.1) sequence (270 bp). The normal and NCBI sequences of the HLA-DRB1*03 exon2 allele were identical at 270 bp. The A1, A2, A3, and A4 groups differed (SNPS) from the NCBI reference and the normal sequence in both forward and reverse directions. First, the forward direction was changed to 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 positions 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 are missing nucleotides in positions 22, 24, 50, and 211, and extra nucleotides at position 196. While in the reverse direction, there were replacements 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 positions 24, 55, 58, 77, 82, 88, 104, 107, 114, 160, 163, 166, 167, and 196, respectively. Also, there is a is a missing nucleotide at position 171 and extra nucleotides at position 186.

The substitution in nucleotides caused a change in arrangement and converted the protein amino acids, which were able to be produced from the HLA-DRB1*03 exon2 allele, to different amino acid sequences, as shown in Tables 1 and 4. With regard to the forward direction, we noticed an identity between the NCBI reference, normal group, group 1, group 2, and group 4 of patients studied. Also, group 3 had tiny variations in comparing it with the NCBI reference and the normal group. These variations involve glycine, proline, threonine, and tryptophan. The reverse direction and variations in the amino acid count and percentage among the groups were also noticed. Although the NCBI reference, the normal group, group 1, group 2, and group 4 were identical, as in the forward direction. Variations were found between Group 3 and the rest of the

of the groups. The amino acids involved in these variations were Alanine, Arginine, Asparagine, Cysteine, Glycine, Histidine, Proline, and Threonine, as shown in tables (2) and (5).

The RNA molecule is an ordered sequence of nucleotides that contain one of the four bases: adenine, cytosine, guanine, and uracil, arranged in the 5' to 3' direction. Pairing (via hydrogen bonds) these four bases within an RNA molecule gives rise to the secondary structure. The canonical base pairs include the Watson-Crick base pairs (A-U and G-C) and the wobble base pair (G-U). These base pairs often result in the formation of a nested structure, in which multiple stacked base pairs form a helix and one or multiple unpaired base pairs form a loop [51] (Zhao et al., 2021). So, modifications occur in the secondary structure in the event of any change in the sequence of RNA. These modifications can affect the free energy values. This is due to the fact that it requires more energy to break the hydrogen bonds. We obtained this secondary structure and the minimum free energy values by using the online RNA Fold Web Server, 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 the consensus RNA sequence transcribed from the HLA-DRB1*03 exon2 allele in the forward direction in groups 2 and 3 T1DM patients studied were slightly lower than that of the NCBI reference, normal group, group 1, and group 4. Also, the RNA secondary structure stability in these studied groups will be different than that of the NCBI reference. and normal sequences. Consequently, the rate of protein translation for group 2 and group 3 studied groups will be slightly greater than that of the NCBI reference. and the normal sequences, but it could be nonfunctional.

On the other hand, in the reverse direction of the HLA-DRB1*03 exon 2 allele sequence, the MFE in group 3 was higher than normal in the NCBI reference sequence, group 1, group 2, and group 4. So, the protein production rate would be less than normal, but it may be non-functional due to the change in amino acid sequence.

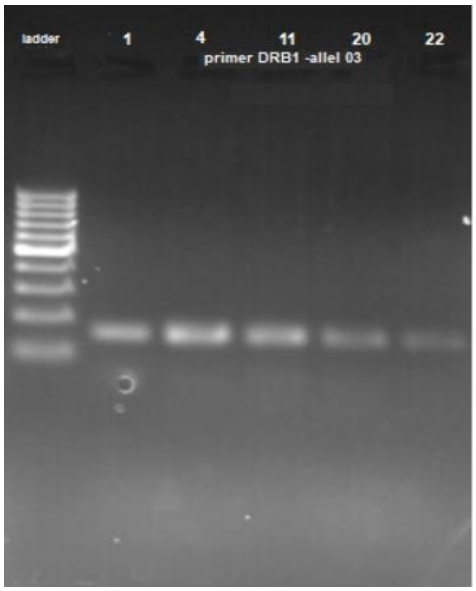


Figure 1: PCR products of the HLA exon2 gene alleles DRB1*03 with approximately 270 bp.

```
A1      CACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATGGGACGGAGCGGG
A2      CACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATGGGACGGAGCGGG
A3      CACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATGGGACGGAGCGGG
A4      CACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATGGGACGGAGCGGG
N       CACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATGGGACGGAGCGGG
ref     CACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATGGGACGGAGCGGG

con     ************************************************************

A1      TGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGTTCGTGCGCTTCGACAGCG
A2      TGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGTTCGTGCGCTTCGACAGCG
A3      TGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGTTCGTGCGCTTCGACAGCG
A4      TGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGTTCGTGCGCTTCGACAGCG
N       TGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGTTCGTGCGCTTCGACAGCG
ref     TGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGTTCGTGCGCTTCGACAGCG

con     ************************************************************

A1      ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGA
A2      ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGA
A3      ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGA
A4      ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGA
N       ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGA
ref     ACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGTACTGGA

con     ************************************************************

A1      ACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCAGAC
A2      ACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACAGCAGAC
A3      ACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCTGGTGGACACCTACTGCAGAC
A4      ACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCAGAC
N       ACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCAGAC
ref     ACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCAGAC

con     ********************************** ******** **** ******

A1      ACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAG
A2      ACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAG
A3      ACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAG
A4      ACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAG
N       ACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAG
ref     ACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAG

con     *******************************************
```

Figure 2: T-coffee alignment of the exon2 DNA consensus sequence of the HLA-DRB1*03 gene in the forward direction for all groups under study.

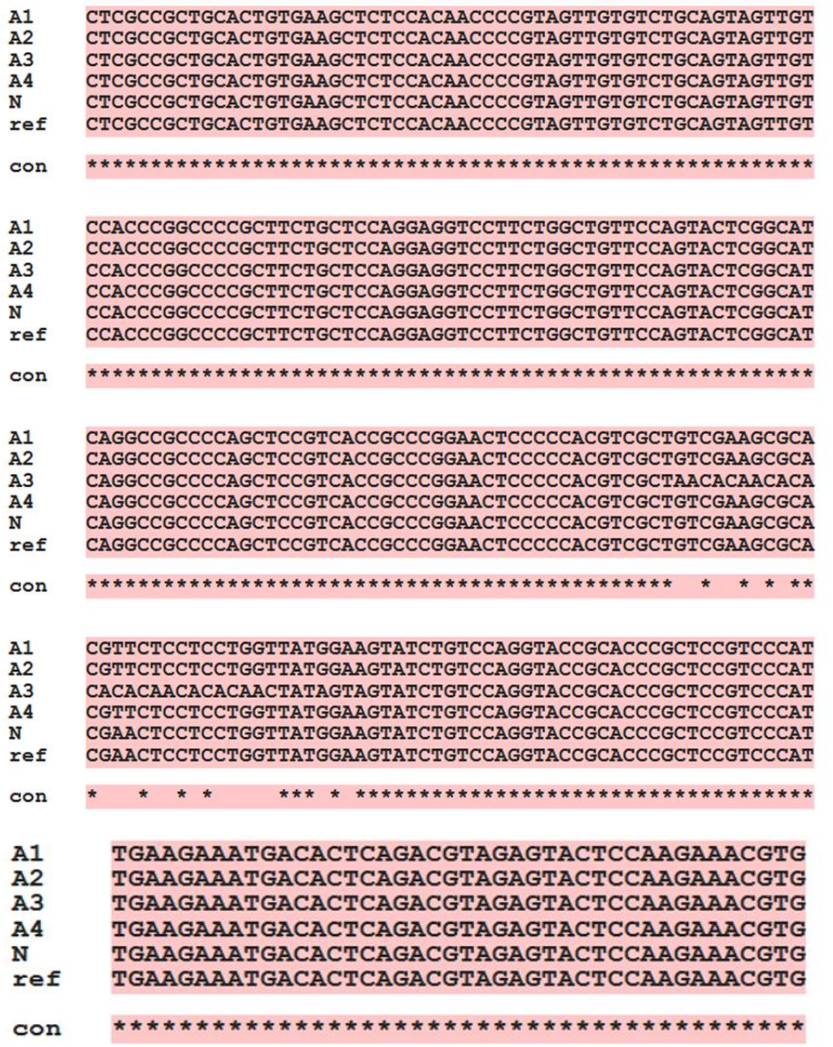


Figure 3: T-coffee alignment of the HLA-DRB1*03 exon2 DNA consensus sequence for all groups under investigation, presented in the reverse orientation.

Figure 4: The encoded amino acids of HLA-DRB1*03 exon2 in the forward orientation across all groups.

Figure 5: The encoded amino acids of HLA-DRB1*03 exon2 in the reverse orientation across all groups.

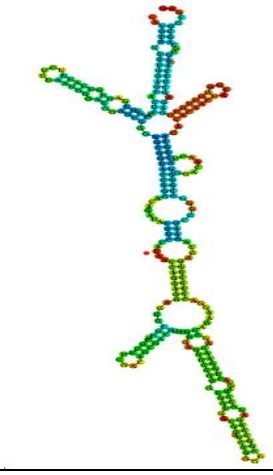
Table 1: The percentage and numbers of amino acids in the protein of HLA-DRB1*03 exon2 in the forward direction for all groups.

	N	N%	A1	A1%	A2	A2%	A3	A3%	A4	A4%
Ala (A)	6	6.742	6	6.742	6	6.742	6	6.742	6	6.742
Arg (R)	10	11.236	10	11.236	10	11.236	10	11.236	10	11.236
Asn (N)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Asp (D)	2	2.247	22.247		2	2.247	2	2.247	2	2.247
Cys (C)	3	3.371	3	3.371	3	3.371	3	3.371	3	3.371
Gln (Q)	1	1.123	1	1.123	1	1.123	1	1.123	1	1.123
Glu (E)	1	1.123	1	1.123	1	1.123	1	1.123	1	1.123
Gly (G)	12	13.483	12	13.483	12	13.483	11	12.359	12	13.483
His (H)	1	1.123	1	1.123	1	1.123	1	1.123	1	1.123
Ile (I)	2	2.247	2	2.247	2	2.247	2	2.247	2	2.247
Leu (L)	4	4.494	4	4.494	4	4.494	4	4.494	4	4.494
Lys (K)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Met (M)	2	2.247	2	2.247	2	2.247	2	2.247	2	2.247
Phe (F)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Pro (P)	1	1.123	1	1.123	1	1.123	2	2.247	1	1.123
Ser (S)	19	21.348	19	21.348	19	21.348	19	21.348	19	21.348
Thr (T)	16	17.977	16	17.977	16	17.977	15	16.854	16	17.977
Trp (W)	7	7.865	7	7.865	7	7.865	8	8.989	7	7.865
Tyr (Y)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Val (V)	2	2.247	2	2.247	2	2.247	2	2.247	2	2.247

Table 2: The percentage and numbers of amino acids in the protein of HLA-DRB1*03 exon2 in the reverse direction for all groups.

	N	N%	A1	A1%	A2	A2%	A3	A3%	A4	A4%
Ala (A)	11	12.644	11	12.644	11	12.644	10	11.628	11	12.644
Arg (R)	8	9.195	8	9.195	8	9.195	6	6.977	8	9.195
Asn (N)	2	2.299	2	2.299	2	2.299	4	4.651	2	2.299
Asp (D)	1	1.149	1	1.149	1	1.149	1	1.163	1	1.149
Cys (C)	6	6.896	6	6.896	6	6.896	5	5.814	6	6.896
Gln (Q)	3	3.448	3	3.448	3	3.448	3	3.488	3	3.448
Glu (E)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Gly (G)	8	9.195	8	9.195	8	9.195	6	6.977	8	9.195
His (H)	2	2.299	2	2.299	2	2.299	5	5.814	2	2.299
Ile (I)	1	1.149	1	1.149	1	1.149	1	1.163	1	1.149
Leu (L)	2	2.299	2	2.299	2	2.299	2	2.326	2	2.299
Lys (K)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Met (M)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Phe (F)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Pro (P)	19	21.839	19	21.839	19	21.839	17	19.767	19	21.839
Ser (S)	14	16.092	14	16.092	14	16.092	14	16.279	14	16.092
Thr (T)	7	8.060	7	8.060	7	8.060	9	10.465	7	8.060
Trp (W)	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Tyr (Y)	1	1.149	1	1.149	1	1.149	1	1.163	1	1.149
Val (V)	2	2.299	2	2.299	2	2.299	2	2.326	2	2.299

Table 3: The predicted secondary structure of HLA-DRB1 *03 allele in the forward direction.

The predicted secondary structure of RNA for normal group, NCBI (HE963767.1) reference, A1, and A4 sequence. The minimum free energy is **-93.02** kcal/mol	

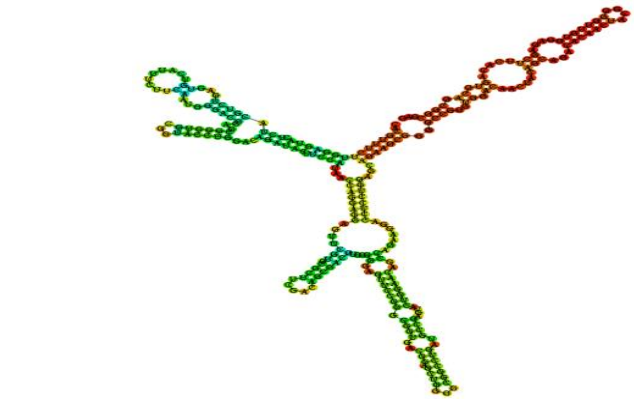
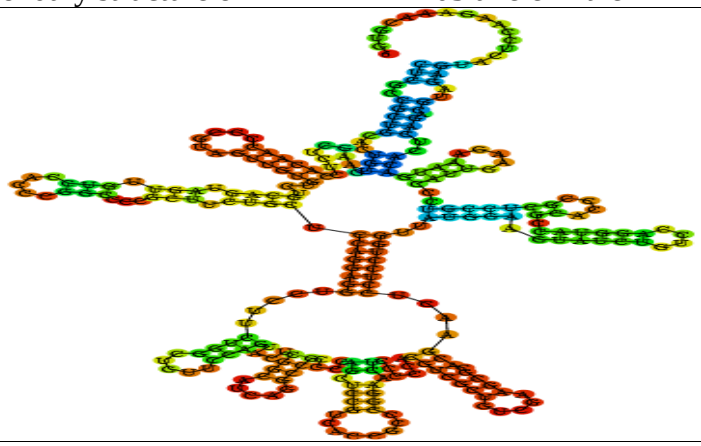
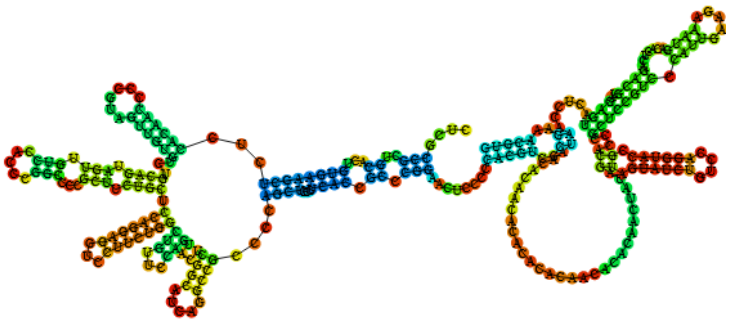
The predicted secondary structure of RNA A2 sequence. The minimum free energy is **-93.42** kcal/mol	
The predicted secondary structure of RNA A3 sequence. The minimum free energy is **-93.22** kcal/mol	

Table 4: The predicted secondary structure of HLA-DRB1 *03 allele in the reverse direction.

The predicted secondary structure of RNA for N, R, A1, A2, andA4 sequence. The minimum free energy is **-72.64** kcal/mol	
The predicted secondary structure of RNA A3 sequence. The minimum free energy is **-60.86** kcal/mol	

4. Discussion

Type 1 diabetes mellitus (T1DM) is an autoimmune disorder that arises from the interaction of hereditary and environmental elements. The majority of persons who acquire Type 1 Diabetes (T1D) possess a Human Leukocyte Antigen (HLA) composition that contributes to 50% of the genetic susceptibility to the disease. In addition to these HLA haplotypes and the insulin area, which play a significant role in the hereditary aspect, genome-wide association studies have discovered several variations in more than 60 non-HLA gene regions that also contribute to the vulnerability to T1DM [52] (Gootjes et al., 2022).Both genome screens and studies have established that T1DM is a genetically polygenic illness, resulting from the interaction of many susceptibility and protective genes[53] (Field et al., 2002). Out of the top 10 nations with the highest prevalence of diabetes worldwide, six of them are located in the Middle East and North Africa. The countries mentioned are Bahrain, Egypt, Kuwait, Oman, Saudi Arabia, and the United Arab Emirates [11] (International Diabetes Federation, 2015). Egypt is the country that has been impacted the most out of all the countries in the Eastern Mediterranean and Middle Eastern regions. It accounts for around 25% of the total number of cases in the region [12] (International Diabetes Federation, 2012). The prevalence of this condition in children under the age of 15 ranges from 1/100 000 per year in Pakistan to 8/100 000 per year in Egypt [13] (Soltèsz et al., 2006).

El-Ziny et al. (2014) conducted the first study on the epidemiology of childhood T1DM in the Nile Delta region of Egypt. They discovered that the incidence and prevalence of T1DM have been increasing over the previous 18 years (1994-2011). According to El-Ziny et al. (2014), females had a higher incidence and prevalence of the condition, and a greater number of cases were reported from rural areas [14].

This latest molecular research identifies differences in the partial HLA-DRB1*03 exon2 allele sequence between patients with T1DM and sequences from the NCBI database. Furthermore, ascertain the mutations that distinguish certain T1DM groups from individuals without diabetes. For this study, the blood tissue of the volunteer was utilized to extract the whole DNA, which was then amplified using the polymerase chain reaction (PCR) method to amplify the specific sequences of interest. The target bands had a length of 270 base pairs in both the forward and reverse directions. Subsequently, it was visualized via electrophoresis, removed

from the gel, and purified. The T-coffee and RNAfold web service, along with other bioinformatics tools, were utilized to assess genetic variations in specific areas of the HLA-DRB1*03 exon2 allele.

This work employed the Sanger DNA sequencing technology to identify the partial sequence of the HLA-DRB1*03 exon 2 alleles. The purpose was to find any mutations that occurred in T1DM patients compared to the normal population and the NCBI reference. Bioinformatics tools are essential for obtaining high-quality discoveries from the sequencing of raw materials. Therefore, these tools are necessary to get the intended conclusion [48] (Detwiler et al.,2004).

This study employed this method to investigate the partial sequence of the HLA-DRB1*03 exon 2 allele in both directions, aiming to identify mutations that occurred in patients with T1DM compared to the general population and the NCBI reference. Bioinformatics tools are essential for obtaining high-quality discoveries from the sequencing of raw materials. Therefore, these tools are necessary to get the intended conclusion [48] (Detwiler et al., 2004).

Koonin asserted that the field of biological sciences can be categorized into three distinct areas [54] (Koonin.,2012). DNA serves as the repository of genetic information, RNA facilitates the conversion of genetic information into proteins, and proteins play a crucial role in numerous biological processes [49] (Notredame et al., 2000). Hence, it was appropriate to employ the Sequence Manipulation Suite to forecast the RNA and the translated protein sequence for subsequent investigation.

Single nucleotide polymorphisms (SNPs) are the most common form of genetic variation and form the basis for the majority of molecular markers. Prior to their utilization in direct sequence-based SNP detection or any related SNP assay, it is necessary for them to be acknowledged. Sequence fragments from different genotypes that come from the same part of the genome can be aligned to detect differences in the sequence. The nucleotides that do not match correspondingly are most likely SNPs or insertions or deletions [55]. (Oeveren & Janssen, 2009).

All patients had a substantial number of missing and substituted nucleotides in exon2 of the HLA-DRB1*03 genotype. These mutations cause alterations in the anticipated amino acid sequences of the HLA-DRB1*03 exon2 allele proteins in the patients. The single-stranded RNA molecules undergo folding by Watson-

Crick base pairing principles, resulting in the formation of secondary structures. These structures consist of a helix formed by several stacked base pairs, together with one or more unpaired base pairs that form a loop [51] (Zhao et al., 2021).

The RNA secondary structure is accountable for various biological tasks and holds a significant position in genome instability, cellular evolution, genetic diversity, mutation, and illness. The citation is from Hurst et al. in 2024[56]. Various methodologies and procedures can be used to estimate the secondary structure and evaluate the minimal free energy (MFE) of RNA. The source of this information is Knudsen and Hein's publication from 2003[57]. There is no text provided. The secondary structure of RNA is crucial for protein biosynthesis. It can hinder translation by impeding the initiation and movement of ribosomes along the mRNA, resulting in reduced protein production. This plays a significant role in regulating gene expression by modifying the minimum energy needed to unravel the secondary structure and initiate translation [58] (Gaspar et al., 2013).

The RNA secondary structure was reconstructed using the RNAFOLD online service in this study. The findings indicated that the minimum free energy (MFE) computed for RNA secondary structures forecasted from the consensus RNA sequence transcribed from the HLA-DRB1*03 exon2 allele in the forward direction in groups 2 and 3 T1DM patients examined were marginally lower than that of the NCBI reference, normal group, group 1, and group 4. Moreover, the stability of the RNA secondary structure in these groups under investigation will vary from that of the NCBI reference and normal sequences. As a result, the rate of protein synthesis for group 2 and group 3 examined groups will be slightly higher than that of the NCBI reference and the usual sequences, although it may not be functional.

Conversely, the HLA-DRB1*03 exon 2 allele sequence exhibited a larger MFE in group 3 compared to the NCBI reference sequence, group 1, group 2, and group 4. Therefore, the pace of protein production would be reduced compared to the usual rate, and the protein may not function properly because of the alteration in the sequence of amino acids.

Overall, this study demonstrated that the HLA-DRB1*03 exon2 allele in the four groups examined exhibited mutations in comparison to the normal and NCBI.

5. Reference

[1] Quattrin, T., Mastrandrea, L. D., & Walker, L. S. (2023). Type 1 diabetes. *The Lancet*, *401*(10394), 2149-2162.

[2] Holt, R. I., DeVries, J. H., Hess-Fischl, A., Hirsch, I. B., Kirkman, M. S., Klupa, T., Ludwig, B., Nørgaard, K., Pettus, J., Renard, E., Skyler, J.S., Snoek, F. J., Weinstock, R. S. & Peters, A. L. (2021). The management of type 1 diabetes in adults. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes care*, *44*(11), 2589-2625.

[3] Stene, L. C., & Tuomilehto, J. (2024). Epidemiology of type 1 diabetes. *Textbook of diabetes*, 41-54.

[4] Katsarou, A., Gudbjörnsdottir, S., Rawshani, A., Dabelea, D., Bonifacio, E., Anderson, B. J., Jacobsen,L. M., Schatz, D. A. & Lernmark, Å. (2017). Type 1 diabetes mellitus. *Nature reviews Disease primers*, *3*(1), 1-17.

[5] Miller, R. G., Secrest, A. M., Sharma, R. K., Songer, T. J., & Orchard, T. J. (2012). Improvements in the life expectancy of type 1 diabetes: the Pittsburgh Epidemiology of Diabetes Complications study cohort. *Diabetes*, *61*(11), 2987-2992.

[6] Esposito, S., Toni, G., Tascini, G., Santi, E., Berioli, M. G., & Principi, N. (2019). Environmental factors associated with type 1 diabetes. *Frontiers in endocrinology*, *10*, 592.

[7] Centers for Disease Control and Prevention. (2017). Rates of new diagnosed cases of type 1 and type 2 diabetes on the rise among children, teens. *Obtained from the world wide web at: https://www. cdc. gov/media/releases/2017/p0412-diabtes-rates. html*.

[8] Mobasseri, M., Shirmohammadi, M., Amiri, T., Vahed, N., Fard, H. H., & Ghojazadeh, M. (2020). Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. *Health promotion perspectives*, *10*(2), 98.

[9] International Diabetes Federation. (2017). IDF Diabetes Atlas. 8th Edition. Brussels: International Diabetes Federation.

[10] Shojaeian, A., & Mehri-Ghahfarrokhi, A. (2018). An overview of the epidemiology of type 1 diabetes mellitus. *Int J Metab Syndr*, *2*(1), 1-4.

[11] International Diabetes Federation. Idf Diabetes Atlas. 2015. Available from: http://www.diabetesat : las.org/resources/2015-atlas.html%5Cnhttps.

[12] International Diabetic Foundation. (2012). Diabetes at a glance.

[13] Soltèsz, G., Patterson, C., & Dahlquist, G. (2006). Global trends in childhood type 1 diabetes. In: Diabetes Atlas, 3rd ed. International Diabetes Federation, Brussels, 2006:153-190.

[14] 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.

[15] Attia, F. A. F., Elfettoh, M. N. A., Alghobashy, A. A., & Fattah, N. R. A. (2023). Brief Overview about Type 1 Diabetes Mellitus among Children. *Tobacco Regulatory Science (TRS)*, 3663-3689.

[16] Redono, M.J. & Fain, M. (2001): Autoimmune and familial risk factors of type 1A diabetes. Recent Horm Res; 56: 69-89.

[17] Redondo, M. J., Jeffrey, J., Fain, P. R., Eisenbarth, G. S., & Orban, T. (2008). Concordance for islet autoimmunity among monozygotic twins. *New England Journal of Medicine*, *359*(26), 2849-2850.

[18] Jones, E.Y.; Fugger, L.; Strominger, J. L., & Siebold, C. (2006): MHC class II proteins and disease: a structural perspective. Nat Rev Immunol.6(4):271–82. doi:10.1038/nri1805.

[19] 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. Am J Hum Genet. 1996;59(5):1134–48.

[20] Morel, P.A., Dorman, J.S., Todd, J.A., McDevitt, H.O., & Trucco, M. (1988). Aspartic acid at position 57 of the HLA-DQ beta chain protects against Type I diabetes: a family study. Proc. Natl. Acad. Sci. U.S.A. 85, 8111–8115.

[21] Todd, J.A., Bell, J.I., & McDevitt, H.O. (1987). HLA-DQb gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus.Nature 329, 599–604.

[22] Noble, J. A., Valdes, A. M., Varney, M. D., Carlson, J. A., Moonsamy, P., Fear, A, L., Lane, J. A., Lavant, E., Rappner, R., Louey,A., Concannon, P., Mychaleckyj, J. C.,& Erlich.(2010) HLA class I and genetic susceptibility to type 1 diabetes: results from the type 1 diabetes genetics consortium. Diabetes.2010;59(11):2972–9. doi:10.2337/db10-0699.

[23] Besser, R. E., Bell, K. J., Couper, J. J., Ziegler, A. G., Wherrett, D. K., Knip, M.,Speake, C., Casteels, K., Driscoll, K. A., Jacobsen, L., Craig, M. E., & Haller, M. J. (2022). ISPAD Clinical Practice Consensus Guidelines 2022: Stages of type 1 diabetes in children and adolescents. *Pediatric diabetes*, *23*(8).

[24] Bonifacio, E., Beyerlein, A., Hippich, M., Winkler, C., Vehik, K., Weedon, M. N., Lamimighofer, M., Hattersley, A. T., Krumiek, J., Frohnert, B. I., Steck, A. K., Hagopian, W. A., Krischer, J.P., Lernmark, A., Rewers, M.J., She, J., Toppari, J., Akolkar, B., Oram, R. A., Rish, S. S., Ziegler, A., & TEDDY Study Group. (2018). Genetic scores to stratify risk of developing multiple islet autoantibodies and type 1 diabetes: a prospective study in children. *PLoS medicine*, *15*(4), e1002548.

[25] Polychronakos,C., & Li ,Q.(2011). Understanding type 1 diabetes through genetics: advances and prospects. Nat Rev Genet. 2011;12(11):

781–92. doi:10.1038/nrg3069.

[26] Noble, J. A., Besançon, S., Sidibé, A. T., Rozemuller, E. H., Rijkers, M., Dadkhodaie, F., Bruin, H. D., Kooij, J., Martin, H. R. N.,Ogle,G. D., & Mack, S. J. (2024). Complete HLA genotyping of type 1 diabetes patients and controls from Mali reveals both expected and novel disease associations. *HLA*, *103*(1), e15319.

[27] Noble, J. A. (2015). Immunogenetics of type 1 diabetes: a comprehensive review. *Journal of autoimmunity*, *64*, 101-112.

[28] Cudworth, A. G.& Woodrow, J. C. (1974). HL-A antigens and diabetes mellitus. The Lancet, 304(7889), 1153.

[29] Nerup, J., Platz, P., Ortved Andersen O, Christy, M., Lyngsøe, J., Poulsen, J. E., Ryder, L.P., Thomsen, M., Staub Nielsen L. & Svejgaard, A. (1974). HL-A antigens and diabetes mellitus. Lancet. 2(7885):864–6.

[30] Singal, D. P., & Blajchman, M. A. (1973). Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. Diabetes. 1973 Jun;22(6):429-32. doi: 10.2337/diab.22.6.429. PMID: 4541338.

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

[32] Noble, J. A., & Valdes, A. M. (2011): Genetics of HLA region in the prediction of type 1 diabetes. Curr Diab Rep.

[33] Varney, M. D., Valdes, A. M., Carlson, J. A., Noble, J. A., Tait, B. D., Bonella, P., Lavant, E., Fear, A. L., Louey, A., Moonsamy, P., Mychaleckyj J. C. (2010). Type 1 Diabetes Genetics Consortium. HLA DPA1, DPB1 alleles and haplotypes contribute to the risk associated with type 1 diabetes: analysis of the type 1 diabetes genetics consortium families. *Diabetes*, *59*(8), 2055-2062.

[34] Howson, J. M., Walker, N. M., Clayton, D. & Todd, J. A. (2009). Confirmation of HLA class II independent type 1 diabetes associations in the major histocompatibility complex including HLA-B and HLA-A. Diabetes Obes Metab 11(Suppl. 1):31–45.

[35] Erlich, H., Valdes, A. M., Noble, J., Carlson, J. A., Varney, M., Concannon, P., Mychaleckyj, J. C., Todd, J. A., Bonella, P., Fear, A. L., Lavant, E., Louey, A. & Moonsamy, P. (2008). Type 1 Diabetes Genetics Consortium. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1diabetes genetics consortium families. Diabetes 57:1084–1092.

[36] Nejentsev, S., Howson, J. M., Walker, N. M., Szeszko, J., Field, S. F., Stevens, H.E., Reynolds, P., Hardy, M., King, E., Masters, J., Hulme, J., Maier, L. M., Smyth, D., Bailey, R., Cooper, J.D., Ribas, G., Campbell, R. D., Clayton, D. G., & Todd, J. A. (2007). Wellcome Trust Case Control Consortium. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature 2007; 450:887–892.

[37] Johansson, S., Lie, B. A., Todd, J. A., Pociot, F., Nerup, J., Cambon-Thomsen, A., Kockum,I., Akselsen, H. E., Thorsby, E., & Undlien, D. E. (2003). Evidence of at least two type 1 diabetes susceptibility genes in the HLA complex distinct from HLA-DQB1,-DQA1 and–DRB1. *Genes & Immunity*, *4*(1), 46-53.

[38] Thorsby, E. (1997).HLA associated diseases. Anniversary review article Hum Immunol 1997; 53: 1–11.

[39] She, J. X. (1996).Susceptibility to type 1 diabetes: HLA-DQ and DR revisited. Immunol Today 1996; 17: 323–329.

[40] Robinson, J., Halliwell, J. A., Hayhurst, J. D., Flicek, P., Parham, P., & Marsh, S. G. (2015). The IPD and IMGT/HLA database: allele variant databases. *Nucleic acids research*, *43*(D1), D423-D431.

[41] González-Galarza, F. F., Takeshita, L. Y. C., Santos, E. J. M., Kempson, F., Maia, M. H. T., Silva, A. L. S. da, Silva, A. L. T., Ghattaoraya, G.G., Alfirevic, A., Jones, A.R. & Middleton, D. (2014). Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Research, 43(D1), D784–D788. doi:10.1093/nar/gku1166.

[42] Moyer, A. M. & Gandhi, M. J. (2022). Human leukocyte antigen (HLA) testing in pharmacogenomics. Methods in Molecular Biology. 2547:21-45.

[43] Bravo-Egana, V., Sanders, H. & Chitnis, N. (2021). New challenges, new opportunities: Next generation sequencing and its place in the advancement of HLA typing. Human Immunology. 82(7):478-487.

[44] Dunn, P. P. (2011). Human leucocyte antigen typing: Techniques and technology, a critical appraisal. International Journal of Immunogenetics. 38(6):463-473.

[45] Rowe, R.E., Leech, N.J., Nepom, G.T., & McCulloch, D.K. (1994). High genetic risk for IDDM in the Pacific Northwest: first report from the Washington State diabetes prediction study. Diabetes 43, 87–94.

[46] Dashti, M., Nizam, R., Jacob, S., Al-Kandari, H., Al Ozairi, E., Thanaraj, T. A., & Al-Mulla, F. (2023). Association between alleles, haplotypes, and amino acid variations in HLA class II genes and type 1 diabetes in Kuwaiti children. *Frontiers in Immunology*, *14*, 1238269.

[47] Wu, F., Zhang, W., Zhang, L., Wu, J., Li, C., Meng, X., Wang, X., He, P., & Zhang, J. (2013). NRAMP1, VDR, HLA-DRB1, and HLA-DQB1 gene polymorphisms in susceptibility to tuberculosis among the Chinese Kazakh population: a case-control study. *BioMed research international*, *2013*.

[48] Detwiler, M., Hamp, T., & Kazim, L. (2004)."DNA sequencing using the liquid polymer POP-7™ on an ABI PRISM® 3100 Genetic Analyzer." Biotechniques. 36(6):932-933.

[49] 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.

[50] 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.

[51] Zhao, Q., Zhao, Z., Fan, X., Yuan, Z., Mao, Q., & Yao, Y. (2021). Review of machine learning methods for RNA secondary structure prediction. *PLoS computational biology*, *17*(8), e1009291.

[52] Gootjes, C., Zwaginga, J. J., Roep, B. O., & Nikolic, T. (2022). Functional impact of risk gene variants on the autoimmune responses in type 1 diabetes. *Frontiers in immunology*, *13*, 886736.

[53] Koonin, E.V. Does the central dogma still stand?. Biol Direct 7, 27 (2012).https://doi.org/10.1186/1745-6150-7-27

[54] Field, L. L., Larsen, Z., Pociot, F., Nerup, J., Tobias, R., & Bonnevie-Nielsen, V. (2002). Evidence for a locus (IDDM16) in the immunoglobulin heavy chain region on chromosome 14q32. 3 producing susceptibility to type 1 diabetes. *Genes & Immunity*, *3*(6), 338-344.

[55] 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.

[56] Hurst, D., Iliopoulos, C. S., Lim, Z., & Moraru, I. (2024, June). IRfold: An RNA Secondary Structure Prediction Approach. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 131-144). Cham: Springer Nature Switzerland.

[57] Knudsen, B., & Hein, J. (2003). Pfold: RNA secondary structure prediction using stochastic context-free grammars. Nucleic acids research, 31 (13), 3423-3428.

[58] 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.