

RESEARCH

Open Access



# Investigation of missense mutation-related type 1 diabetes mellitus through integrating genomic databases and bioinformatic approach

Dyonisa Nasirochmi Pakha<sup>1</sup> , Ratih Dewi Yudhani<sup>1\*</sup> and Lalu Muhammad Irham<sup>2</sup>

## Abstract

Though genes are already known to be responsible for type 1 diabetes mellitus (T1DM), the knowledge of missense mutation of that disease gene has still to be under covered. A genomic database and a bioinformatics-based approach are integrated in the present study in order to address this issue. Initially, nine variants associated with T1DM were retrieved from the GWAS catalogue. Different genomic algorithms such as PolyPhen2.0, SNPs and GTEx analyser programs were used to study the structural and functional effects of these mutations. Subsequently, SNPnexus was also employed to understand the effect of these mutations on the function of the expressed protein. Nine missense variants of T1DM were identified using the GWAS catalogue database. Among these nine SNPs, three were predicted to be related to the progression of T1DM disease by affecting the protein level. *TYK2* gene variants with SNP rs34536443 were thought to have a probably damaging effect. Meanwhile, both *COL4A3* and *IFIH1* genes with SNPs rs55703767 and rs35667974, respectively, might alter protein function through a possibly damaging prediction. Among the variants of the three genes, the *TYK2* gene with SNP rs34536443 had the strongest contribution in affecting the development of T1DM, with a score of 0.999. We sincerely hope that the results could be of immense importance in understanding the genetic basis of T1DM.

**Keywords** Bioinformatics, Genomic database, Genomic variant, T1DM

## 1 Introduction

Type 1 diabetes mellitus (T1DM) is a chronic autoimmune disease marked by elevated blood glucose levels (hyperglycaemia) due to insulin deficiency. This deficiency is caused by the destruction of pancreatic islet  $\beta$ -cells, resulting from autoimmunity (autoimmune T1DM) and, in rare cases, there was a contribution of strong genetic factor (idiopathic T1DM) [1, 2]. Unlike type 2 diabetes mellitus (T2DM), the population of

T1DM is only 10–15% among all individuals with diabetes and is commonly found at an early age, below 15 years old [1]. However, the incidence of T1DM is expected to continue to rise, with approximately 90,000 children diagnosed annually [3].

Research predominantly centres on T2DM due to its prevalence, whilst studies on T1DM remain limited. T1DM necessitates particular concern due to some challenges, including misdiagnosis, underdiagnosis, high-risk complications, and premature mortality [4]. According to Lind et al. [5], patients with T1DM had a higher risk of death from any cause, primarily cardiovascular disease, which is more than twofold compared to the general population. Moreover, this risk increases in patients with poor hyperglycaemic control. Hence, this disease

\*Correspondence:

Ratih Dewi Yudhani

ratihyudhani@staff.uns.ac.id

<sup>1</sup> Department of Pharmacology, Faculty of Medicine, Universitas Sebelas Maret, Surakarta 57126, Indonesia

<sup>2</sup> Faculty of Pharmacy, Universitas Ahmad Dahlan, Yogyakarta 55166, Indonesia



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

continues as one of the major challenges for clinicians and researchers.

Managing T1DM mostly depends on insulin to prevent severe illness and death as well as to lower the chances of long-term macrovascular and microvascular complications [4]. Despite the presence of advanced insulins, T1DM individuals are still at a high risk of severe complications. Novel approaches are required for the prevention and treatment of T1DM [6]. Autoimmunity is a key feature of T1DM; thus, an intervention that alters the immune system could be a promising approach for treating the illness. This discovery indicates that precision medicine focusing on causative genes and the pathophysiology of complex diseases like T1DM, when fully understood, could offer more effective management [7].

The pathogenesis of T1DM involves a complex interaction between autoantibodies, genetics, and environmental factors [1]. The early autoantibodies detected in T1DM are insulin or anti-glutamic acid decarboxylase (GAD) autoantibodies. Other antibodies that can develop are protein tyrosine phosphate-related islet antigen 2 (IA-2) and zinc transporter 8 (ZnT8) autoantibodies in later stages [1, 8]. The development of these autoantibodies is related to genetic differences, age, and environmental exposure [1]. It has been documented that T1DM involves multiple gene abnormalities at different loci [8]. However, genetic risk factors alone are insufficient, with only 30% of the concordance rate of monozygotic twins having T1DM, despite long-term follow-up implicating a higher percentage. *HLA-DR3-DQ2* and *HLA-DR4-DQ8* haplotypes are the major genetic risk factors in the development of  $\beta$ -cell-targeting autoantibodies [9]. As a result of its association with  $\beta$ -cell-targeting autoantibodies, *HLA*-associated risk factors might elevate the risk of T1DM development. Furthermore, various environmental factors, such as viral infections, the initiation of food exposure, and gestational events, might cause the emergence of autoantibodies [1].

It has been documented that genome-wide association studies (GWAS) have recorded over 60 genetic factors that contributed to T1DM risk, marked by single nucleotide polymorphisms (SNPs) [10]. A study by Nyaga et al. [10] has identified T1DM-associated SNPs in regulatory networks, which are associated with the inflammation and destruction of pancreatic  $\beta$ -cells, the signalling of adaptive immune, and the proliferation and activation of immune cells. Although the genomic variants related to T1DM have been largely identified through the GWAS-based approach, the pathogenic variants with missense mutation are still limited to explore. Moreover, the molecular mechanism of some SNPs regions is still unknown. Hence, identifying these genetic factors at an early stage may allow for more time for prevention

or treatment, as well as a slower progression of the disease [8]. This study aims to identify the genetic variants of T1DM associated with SNPs focusing on missense variants by using genomic database and bioinformatic approaches. Hence, this finding may provide an overview of potential biomarker candidates that might contribute to understanding the pathogenesis of T1DM and proposing strategies for future therapeutic approaches.

## 2 Methods

### 2.1 Retrieval of GWAS datasets

The National Human Genome Research Institute (NHGRI) GWAS Catalogue Database (<https://www.ebi.ac.uk/gwas/>) was used to identify the T1DM associated with SNPs [11]. Using the keyword “type 1 diabetes mellitus” (MONDO\_0005147) and including the background as well as child traits data, all available data of associations were downloaded on 16 Jan 2023, resulting in 815 variant and risk alleles. Next, the data was sorted by focusing on the missense variant, odds ratio (OR) > 1, and  $p$ -value <  $10^{-8}$ . The missense variants are genetic variants changing the amino acid sequence of the protein. By this alteration, the pathogenic missense variant interferes with the protein function and affects the phenotypes, whilst the benign missense variants have limited impact [12]. Hence, it is essential to focus on which missense variant is thought to modify the protein function. Meanwhile, OR greater than one is considered to increase the odds among those who are exposed compared to the unexposed [13]. In addition, a  $p$ -value with a threshold below  $10^{-8}$  was applied to distinguish true and false positives [14–16]. This is reasonable since multiple testing corrections were performed in GWAS that may result in a high number of false positive if the significance criterion was 0.05. Hence, using the present threshold ( $p$ -value <  $10^{-8}$ ) improves the robustness and reproducibility of alleged associations [17].

### 2.2 Projection of non-synonymous coding SNPs on protein function

After sorting and removing duplicates using Microsoft Excel, 9 SNPs were obtained. Then, those SNPs were assessed through SNPnexus (<https://www.snp-nexus.org>) by submitting batch queries of the 9 SNPs by selecting related annotation categories, which were PolyPhen Database and 1000 Genome Population Data. The benefits of using SNPnexus are a user-friendly interface, broad database of annotation fields, which accommodate batch searches, data visualization, and do not require substantial programming expertise nor computing resources from users. Thus, this assures that it is still applicable for analysing and interpreting sequence variants in a wide range of biological applications [18].

The PolyPhen Database aided the identification of SNP variants in affecting the protein changes in the disease, which is divided into benign, possibly damaging, and probably damaging [18–22]. This classification is based on the position-specific independent count (PSIC) scores difference between two variants (wild and mutant amino acid) [23]. The prediction of benign (range 0.0–0.49) indicates that the query substitution is likely to be benign with high confidence, whereas probably damaging (range 0.9 to 1) indicates that the query substitution is projected to be damaging with high confidence. As for the prediction of possibly damaging (range 0.5 to 0.89) reflects that the query substitution is expected to be damaging, albeit with low confidence. Hence, this identification would estimate the effects of single amino acid substitution on protein function and structure [23, 24]. The data were extracted on 22 Jan 2023, and finally, after the whole process of analysing, this study revealed 3 SNPs that were projected to be possibly and probably damaging.

### 2.3 Identification of population

The SNPnexus, also, was used to identify the population data of those SNPs using the 1000 Genome Database. The data classified the populations into 5 major continents, namely Africa, America, Europe, East Asia, and South Asia. This database allowed for a thorough description and reference of variation in human genetics; thus, helped to estimate continent-specific allele frequencies [25, 26].

### 2.4 Assessing the distribution of gene expression

The three genes representing the three SNPs were then analysed to review the gene expression in various tissues using GTEx Portal (<http://www.gtexportal.org/home/>), extracted on 25 Jan 2023. The Genotype-Tissue Expression (GTEx) project is a large-scale resource that aids in understanding the complicated patterns of genetic variation and gene regulation found in various human tissue types. This includes analysing and interpreting the GWAS database for translation research. Thus, utilising the GTEx Portal aided in a comprehensive interpretation of this data in a variety of tissues that might be relevant to numerous diseases [27]. This study assessed the bulk tissue gene expression of each gene and focused on the top ten expressions sorted by log scale and median. The whole current method also was applied in several bioinformatics studies including chickenpox disease, systemic lupus erythematosus, and Sjogren’s syndrome [14, 16, 28].

## 3 Results

### 3.1 Screening of T1DM-associated SNPs

In this step, we identified the variants associated with T1DM using a bioinformatics-based approach. It can be seen from the data in Table 1 that nine missense variants of T1DM were identified using the GWAS catalogue database. Among these nine SNPs, three were predicted to be related to the progression of T1DM disease by affecting the protein level. Data in Table 2 depicted that *TYK2* gene variants with SNP rs34536443 were thought to have a probably damaging effect. Meanwhile, both *COL4A3* with SNPs rs55703767 and *IFIH1* genes with rs35667974 might alter protein function through a possibly damaging prediction. Among them, the *TYK2* gene with SNP rs34536443 had the strongest contribution in affecting the development of T1DM, with a score of 0.999.

### 3.2 The distribution of pathogenic variants of T1DM

Table 3 illustrates the proportion of frequencies of three SNPs based on five areas using the 1000 Genome Database. Table 3 shows that rs55703767 was distributed in all populations compared to other SNPs. The highest proportion was found in the European population with the variant rs55703767 (21.17%). In contrast, the allele frequency of SNPs rs34536443 and rs35667974 was below 1% in Africa and South Asia. Meanwhile, both previous SNPs were not shown in the East Asian population (Fig. 1).

### 3.3 The tissue gene expression of T1DM-associated with SNPs

GTEx Portal was used to understand the gene expression of the three genes in various tissues, presented in Figs. 2, 3 and 4. *TYK2* and *IFIH1* genes were highly expressed in cells-EBV-transformed lymphocytes and spleen (Figs. 2 and 3). Both genes were distributed higher in the lung than in other tissues. On the contrary, as shown in Fig. 4,

**Table 1** Missense variant of T1DM associated with SNPs with odds ratio > 1 and *p*-value < 10<sup>-8</sup>

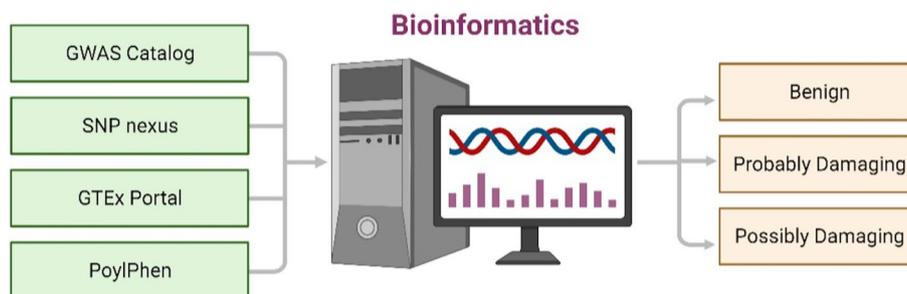
| SNPs       | <i>p</i> -value        | Odds ratio |
|------------|------------------------|------------|
| rs1990760  | 2 × 10 <sup>-14</sup>  | 1.2        |
| rs2304256  | 4 × 10 <sup>-9</sup>   | 1.16       |
| rs2476601  | 2 × 10 <sup>-111</sup> | 2          |
| rs3184504  | 2 × 10 <sup>-38</sup>  | 1.3        |
| rs34536443 | 4 × 10 <sup>-15</sup>  | 1.49       |
| rs35667974 | 9 × 10 <sup>-9</sup>   | 1.69       |
| rs55703767 | 5 × 10 <sup>-12</sup>  | 1.27       |
| rs6043409  | 3 × 10 <sup>-10</sup>  | 1.14       |
| rs763361   | 1 × 10 <sup>-9</sup>   | 1.12       |

**Table 2** Genetic variants related to T1DM associated with SNPs and the effect on protein level

| SNPs       | Chromosome | Variant | Gene   | AA position | Score | Prediction        |
|------------|------------|---------|--------|-------------|-------|-------------------|
| rs34536443 | chr19      | G/C     | TYK2   | 18          | 0.999 | Probably damaging |
|            |            |         |        | 1104        | 0.973 | Probably damaging |
|            |            |         |        | 919         | 0.973 | Probably damaging |
| rs35667974 | chr2       | T/C     | IFIH1  | 884         | 0.999 | Probably damaging |
|            |            |         |        | 923         | 0.859 | Possibly damaging |
| rs55703767 | chr2       | G/A     | COL4A3 | 326         | 0.751 | Possibly damaging |

**Table 3** The distribution of allele frequencies of three SNPs based on the region

| SNPs       | REF allele | Allele<br>ALT allele | Allele frequency |          |            |          |             |
|------------|------------|----------------------|------------------|----------|------------|----------|-------------|
|            |            |                      | African          | American | East Asian | European | South Asian |
| rs34536443 | G          | C                    | 0.0015           | 0.0202   | None       | 0.0288   | 0.0061      |
| rs35667974 | T          | C                    | 0.0008           | None     | None       | 0.0099   | 0.0010      |
| rs55703767 | G          | T                    | 0.0189           | 0.1542   | 0.1290     | 0.2117   | 0.1074      |



**Fig. 1** A schematic model illustrates how an integrated genomic database and bioinformatics approach can be used to identify pathogenic variants for T1DM

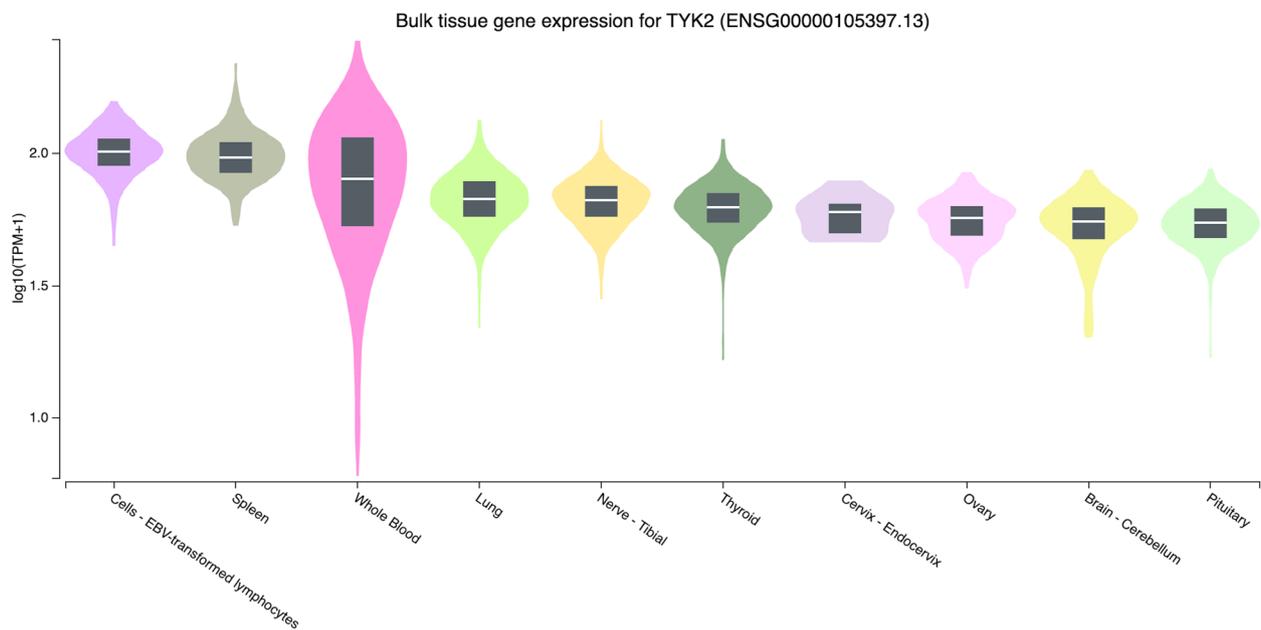
the expression of *COL4A3* was mainly found in the thyroid, pituitary, and kidney. This study highlights that the pathogenic variants of T1DM can be identified by integrating the genomic database and bioinformatics-based approach.

## 4 Discussion

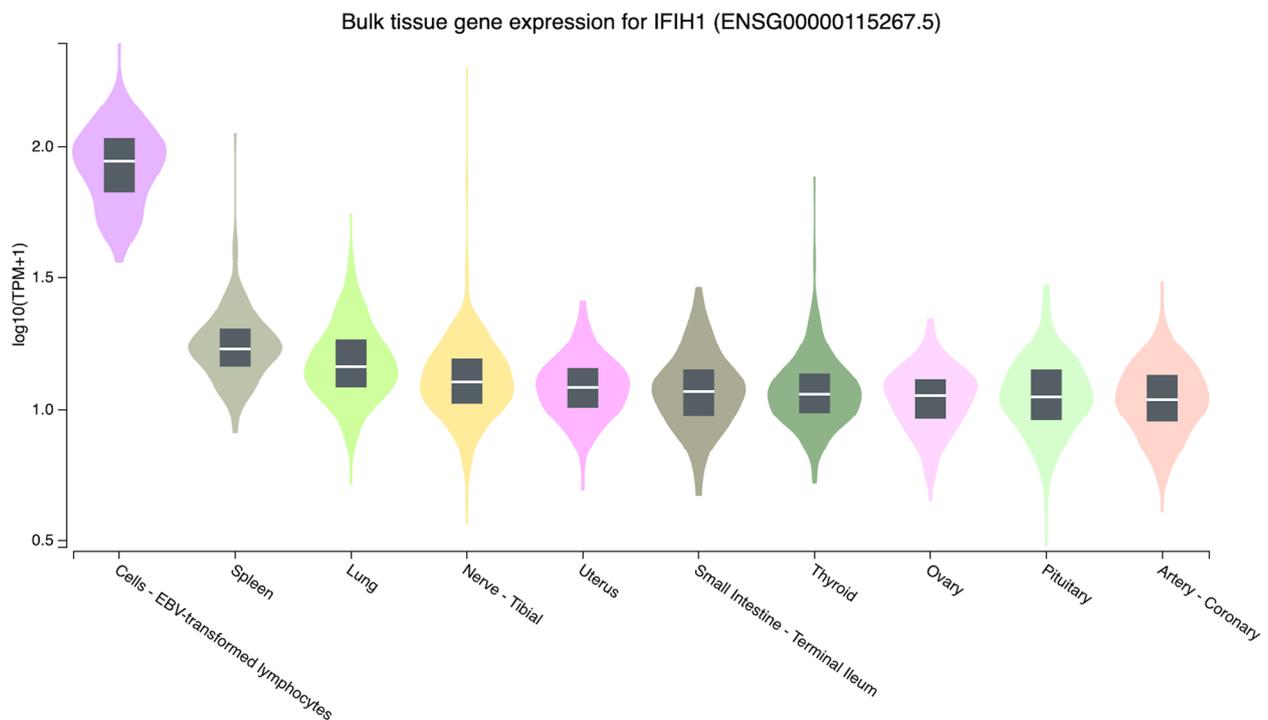
### 4.1 Variants of T1DM-associated SNPs

In this study, we used genomic databases and bioinformatics approaches to identify potential biomarkers for T1DM. interestingly, we identified that the *TYK2* gene variants with SNP rs34536443 were thought to have a probably damaging effect. It has been documented that tyrosine kinase 2 (*TYK2*) belongs to the family of Janus kinase (*JAK*), which is involved in intracellular signalling of cytokines and type I interferons (*IFN-1*) by phosphorylating and activating signal transducers and activators of transcription (*STATs*) [29]. Surprisingly, the *TYK2* contributes to the pathogenesis of T1DM by

enhancing antigen presentation via the upregulation of MHC class I and *CXCL10* chemokine expression. As a result, T-cells will be activated and recruited to pancreatic islets [30]. MHC class I and *CXCL10* chemokine upregulation are seen in islets from patients with T1DM. The downregulation of *TYK2* would reduce *STAT1* phosphorylation and later diminish MHC class I antigen levels in haematopoietic cell lines [31]. In addition, the neutralisation of *CXCL10* chemokine would suppress the emergence of T1DM in nonobese diabetic mice [32]. In line with previous studies, the inhibition of *TYK2* in mature islet cells effectively hampered the *IFNα*-induced MHC Class I upregulation, which led to a significant reduction of T-cell cytotoxicity [33]. Furthermore, the inhibition of *TYK2* in knock-out stem cells (*KO SC*)-islet models affected the endocrine precursor formation during the beginning of islet development, although it did not affect  $\beta$ -cells maturation and function [33]. Therefore, targeting *TYK2* through inhibition might be a promising



**Fig. 2** The ten most expressed tissue genes of TYK2 (based on GTEx Portal)

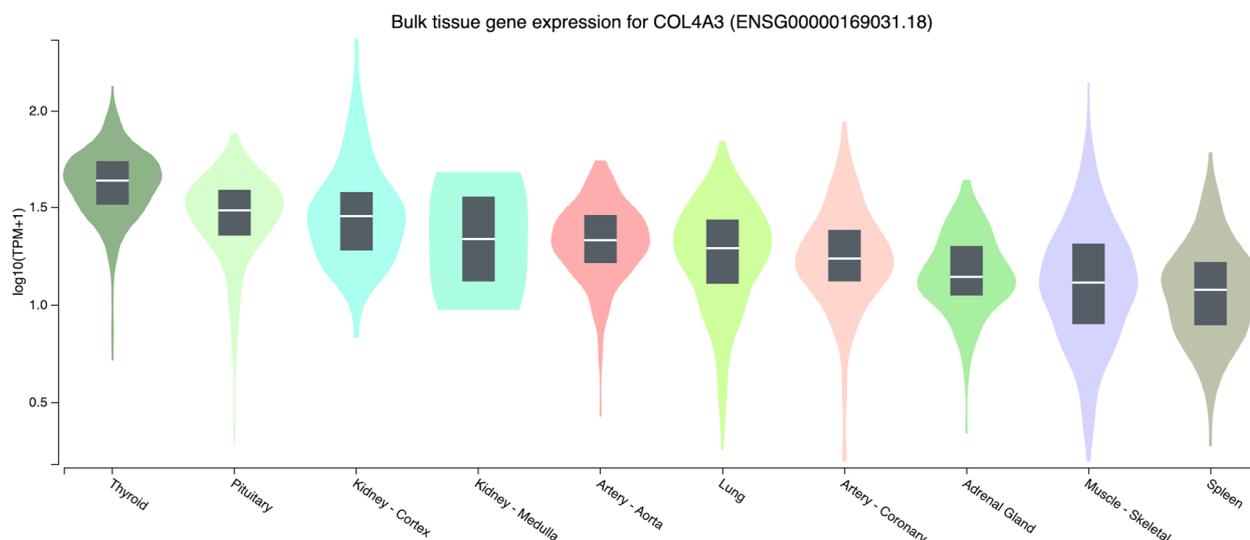


**Fig. 3** The ten most expressed tissue genes of IFIH1 (based on GTEx Portal)

therapeutic approach in T1DM, including preventing the development of T1DM [29].

Moreover, the loss-of-function of *TYK2* with SNPs rs34536443 has been known for the protective function against several autoimmune diseases, including T1DM, by

diminishing IFN-1, IL12, and IL23 signalling [33, 34]. The minor C allele is thought to act as protection for T1DM by disrupting the activity of *TYK2* [30]. In response to type I IFN in vitro, the protective allele of rs34536443 in peripheral blood mononuclear has demonstrated reduced



**Fig. 4** The ten most expressed tissue genes of COL4A3 (based on GTEx Portal)

phosphorylation of TYK2, STAT1, and STAT2 [34]. Other studies revealed that there are impairments in IL-12 and IL23 signalling due to the impact of rs34536443, affecting a low level of pSTAT4 and pSTAT3, leading to reduced diabetogenic Th1 and Th17 populations [30, 34]. Finally, rs34536443 stands out as the only SNP with a noticeable effect on TYK2. As a result, TYK2 may be a promising target for drug-dependent inhibition in various common autoimmune disorders, including T1DM [34].

*Other variants* with SNP rs35667974 encoded the *IFIH1* gene, located at chromosome 2 with the substitution position at 884 and 923, had a contribution to the risk of T1DM and were categorised as probably damaging (score 0.999) and possibly damaging (score 0.859), respectively. Interferon gene induced by helicase C domain 1 (*IFIH1*), also known as melanoma differentiation-associated 5 (MDA-5). *IFIH1* is located on chromosome 2q24.3 and regulates the expression of an early  $\beta$ -responsive type I interferon (IFN) gene. This gene encodes a viral RNA-activated apoptotic protein, thought to play a role in recognising and promoting a clearance response in virus-infected cells [35]. Previous study indicates that a decreased level of expression or functioning of *IFIH1*, known as a viral RNA receptor, protects against T1DM. In T1DM susceptible individuals, both normal or activated antiviral responses could trigger apoptosis of infected pancreatic  $\beta$  cells, which express an elevated quantity of *IFIH1* RNA, and lead to type 1 interferon signalling, improving HLA class I expression on  $\beta$ -cells, thus enhancing cytotoxic CD-8 T cell-mediated destruction [36].

Viruses are shown to play a significant role in initiating the autoimmunity that contributes to the

occurrence of T1DM. Among the multiple viral genotypes studied currently, enteroviruses have been consistently linked to T1DM in humans since EV showed tropism specific to the pancreas [37]. Enteroviruses (EV), also known as small non-envelope RNA viruses, have been revealed to play a crucial role in initiating an autoimmune process, leading to the destruction of the  $\beta$ -cells pancreas. Moreover, the EV genome has been detected in the circulation of T1DM patients [38, 39]. Other evidence indicates that EV infection is present in the  $\beta$ -cells of patients with fulminant diabetes [40, 41]. These underlying *IFIH1* gene polymorphisms can mediate the molecular relationship between specific virus triggers and autoimmune responses in T1DM [42].

A number of *IFIH1* polymorphisms were linked with T1DM, especially the rs1990760 (G/A) polymorphism in the United Kingdom population (substitution of an alanine for a valine in codon 946 of exon 15) being most closely linked with protection from disease development (OR=0.86,  $p=1.42 \times 10^{-10}$  for the G allele) [42]. In contrast, the A allele of the rs1990760 *IFIH1* polymorphism strongly correlated with T1DM in Hungarians (OR=1.29;  $p=0.002$ ) [43]. In a Russian population, two rare *IFIH1* variants, rs35744605 and rs667974, were significantly associated with a reduced incidence of T1DM [44]. The *IFIH1* gene with SNP rs35667974 (substitution isoleucine for valine in codon 923 of exon 14) also showed a protective effect against T1DM (OR=0.43, 95% CI=0.43–0.61,  $p=1.3 \times 10^{-14}$ ) [45].

The *IFIH1* gene significantly affects the innate immune response to viral infection. Binding viral replication-derived dsRNA to *IFIH1* causes immune cells to release proinflammatory cytokines. This local

inflammation and activation of antiviral defence mechanisms are designed to eliminate infection and induce apoptosis in virus-infected cells. Meanwhile, this immune system malfunctions among specific genetically susceptible individuals, eliciting excessive, progressive inflammation and prolonged  $\beta$ -cell death, consequently predisposing them to T1DM. It provides evidence that the *IFIH1* gene is an excellent candidate gene for further investigation into T1DM [37].

Table 2 documented that the *COL4A3* gene with SNP rs55703767, located at chromosome 2 with the substitution position at 329 (G/A), took into account the risk of T1DM and was categorised as possibly damaging (score 0.751). The rs55703767 is a common missense mutation (G > T; Asp326Tyr) in exon 17 of the collagen type IV alpha 3 chain (*COL4A3*) gene. It has been documented that rs55703767 was significantly related to the protection from diabetic nephropathy, any albuminuria, combined phenotype of chronic kidney disease and diabetic nephropathy, and macroalbuminuria. Also, this SNP in *COL4A3* was more strongly associated with men than women [46].

The *COL4A3* gene encodes  $\alpha 3$  chains of type 4 collagen, which is the major structural protein of the kidney basement membranes (BMs). Variants and mutations in the *COL4A3* gene may result in diabetic kidney disease (DKD) in young adults with maturity-onset diabetes, and pathogenic *COL4A3* mutations have been identified for Alport syndrome (AS), a progressive inherited nephropathy [47]. In addition, mutations in *COL4A3* have been reported in patients with focal segmental glomerulosclerosis, a condition in which scar tissue progresses on the glomeruli, which might result in proteinuria and renal failure [46]. Type IV collagen is typically an extracellular structural protein that forms a collagen branch network, a crucial component of BMs [48]. Abnormal collagen homopolymers might result in molecular folding, secretion, and extracellular matrix formation. Misfolded proteins can be secreted into the BM or accumulate in podocytes, altering the glomerular selective barrier structure and initiating downstream pathological pathways [47]. Moreover, dysregulation of interactions between cells, collagen IV basement membrane, cell adhesion, proliferation, survival, and differentiation have been linked to several pathologic illnesses, including chronic kidney disease [46].

T1DM is associated with a lengthy onset of nephropathy. Initially, the patient exhibits hyperfiltration, characterised by high glomerular filtration rate values, about twice as much as the normal value, and occasional episodes of microalbuminuria. Diabetic nephropathy is identified by glomerular hypertrophy and the thickening

of basement, tubular, and glomerular membranes. Additionally, the extracellular matrix accumulation in glomerular membranes was also profound. These pathogenic structures ultimately result in tubulointerstitial and glomerular fibrosis and sclerosis [49]. Multiple mechanisms are involved in the onset and progression of diabetic nephropathy, including the interaction between hyperglycaemia-induced metabolic and hemodynamic alterations and genetic predisposition, thereby establishing the stage for kidney injury [49]. The most significant risk factors for the emergence of diabetic nephropathy are hyperglycaemia, high blood pressure, and genetic predispositions [50]. Therefore, along with the increasing cases of diabetes, diabetic nephropathy is the most common factor causing end-stage renal failure [49]. Based on these, a missense variant in *COL4A3* (rs55703767) was associated with DKD, and *COL4A3* mutations have been implicated in basement membrane diseases such as familial FSGS and Alport symport. Therefore, this risk allele may contribute to the aetiology of diabetic kidney disease, making it a great candidate for target therapy.

Differing from T1DM, neither study up to this current research has identified a connection between variants of *TYK2* and *COL4A3* genes and the likelihood of developing gestational diabetes. Unlike the two genes studied earlier, Kochenborger et al. [51] reported that *IFIH1* gene expression was inversely correlated to HbA1c levels. It indicated a connection between variants of the *IFIH1* gene and its protective effect on the likelihood of developing gestational diabetes mellitus (GDM). Meanwhile, in T2DM, the *TYK2* promoter variant was linked to an increased incidence of diabetes and associated with dysfunctional insulin production [52, 53], whereas the T allele of *COL4A3* was associated with T2DM with a protective role [54]. Most studies have shown that *IFIH1* with rs35667974 protects against the onset of T1DM, but the link with T2DM is unknown and requires further research. However, this differs from *IFIH1* with rs1990760, which demonstrated a lower incidence of T2DM in the Iraqi population [55, 56].

Genome-wide association studies (GWAS) have found a link between genetic differences in single nucleotide polymorphisms (SNPs) and a higher risk of both T2DM and GDM in different groups of people, including polymorphisms at the transcription factor 7-like 2 gene (*TCF7L2*), adiponectin gene (*ADIPOQ*), and fat mass and obesity-associated gene (*FTO*). The relationship has been suggested due to the shared pathophysiological processes of insulin resistance and chronic inflammation in both GDM and T2DM [57]. Hence, to understand the association between each gene in this study with T2DM and GDM, further studies need to be done to validate the association.

#### 4.2 The distribution of pathogenic variants of T1DM-associated with SNPs

This study displayed that Europeans had the highest distribution of the three SNPs related to T1DM rather than other populations. Based on the DIAMOND project, initiated by WHO in 1990, populations from China and South America had the lowest incidence, which was <1/100,000 per year, whilst the highest one was noticed in Sardinia, Finland, Sweden, Norway, Portugal, UK, Canada, and New Zealand [58, 59].

Based on Table 3, the American also had a higher distributed SNP, except for SNP rs35667974. Both rs35667974 and rs34536443 alleles were below 1% in African and South Asian, and those SNPs were not found in the East Asia population. The SEARCH project in the United States reported that the incidence and prevalence rates of T1DM in the youth were higher in non-Hispanic whites (2.0/1000 and 23.6/100,000, respectively) than in other ethnicities [59, 60]. Additionally, according to a meta-analysis by Mobasser et al. [61], the incidence of T1DM in Africa, Asia, Europe, and America was statistically significant, with 8/100,000, 15/100,000, 15/100,000, and 20/100,000 population, respectively. Although the prevalence of T1DM was not significant in Africa at 3.5/10,000, the rates in Asia, Europe, and America were statistically significant, with 6.9/10,000, 12.2/10,000, and 12.2/10,000 people, respectively. Hence, the incidence and prevalence of T1DM are increasing worldwide [61].

Furthermore, the rs55703767 were widely distributed in the five populations compared to the other alleles (Table 3). This SNP has been known as one of the loci associated with DKD and has the strongest signal compared to other loci associated with DKD. The rs55703767 is a common missense variant in exon 17 of *COL4A3*. This variant's risk allele impacted the risk of macroalbuminuria or end-stage renal disease, and its frequency was prevalent among Europeans (20%) and East Asians (13%), but modest in Africans (2%) [62]. This distribution supported the findings in this study, which displayed 21.17% in Europe, 12.9% in East Asia, and 1.89% in Africa (Table 3). Mutations and variants of *COL4A3* contribute to Alport syndrome and DKD [46]. Interestingly, around 40% of those with T1DM have an increased risk of DKD. Individuals with DKD are at a higher risk of developing cardiovascular disease and premature mortality [63].

#### 4.3 The tissue gene expression of T1DM-associated with SNPs

Transcriptional changes in diverse tissues and organs are predominantly responsible for and reflective of tissue and organ dysfunction. In a significant part, transcriptional variation mediates causal links between genotype and complex traits. Thus, insights into tissue-specific

gene expression (TSGE) profiles can contribute to a better understanding of the aetiology of disease [64]. We utilised the Genotype-Tissue Expression (GTEx) portal database at <http://www.gtexportal.org/home/> for determining gene expression levels in human tissues of the genetic variation linked to T1DM. The GTEx is an established gene and its associated tissues database for studying the relationship between genetic variation and gene expression, as well as other molecular phenotypes in multiple reference tissues. It is also beneficial for elucidating the intricate patterns of genetic variation and gene regulation throughout various types of human tissue [65].

The GTEx database documented that the ten most tissues-specific expressions of the *TYK2* and *IFIH1* genes were similar. Both genes were highly expressed in cells-EBV-transformed lymphocytes, spleen, and lung (Figs. 2 and 3). Specifically, *TYK2* was also highly expressed in whole blood. Figure 2 correlates with GenCards' database, showing that *TYK2* is predominantly expressed in spleen and blood cell components such as peripheral blood mononuclear cells (PBMC), B lymphocytes, and CD4 and CD8 T cells [66]. The high level of *TYK2* expression in those tissues might be due to its role in immunity and cytokine signalling mechanisms. Poelzl et al. [67] documented that *TYK2* regulates cytokine signalling and promotes immunity against viral and bacterial infections in humans and mice. In comparison, the *IFIH1* protein is expressed in a tissue-specific manner, including the lung and spleen, according to UniProt (<https://www.uniprot.org/uniprotkb/Q8R5F7/entry>) [68]. This conforms to the depiction in Fig. 3. The spleen contains the three main types of mononuclear phagocytes: macrophages, dendritic cells, and monocytes. They identify pathogens and cellular stress, eliminate dying cells and foreign substances, control tissue homeostasis and inflammatory responses, and contribute to adaptive immunity [69].

In contrast to the former genes, the expression of *COL4A3* was primarily found in the thyroid, pituitary, and kidney (Fig. 4). Similarly, with the HPA dataset, *COL4A3* was primarily expressed in the kidney and thyroid gland. Also, based on the dataset, the lung was in the top three of the expression of *COL4A3* in organs [70]. However, this was slightly different from the data from GTExPortal (Fig. 4), although the lung was still in the top ten of most expressed tissue genes of *COL4A3*. It has been documented that *COL4A3* had a strong signal related to a thinner glomerular basement membrane and protection against albuminuria and diabetic kidney disease [71]. Hence, this might explain the expression in the kidney.

This bioinformatics approach has identified three risk genes associated with T1DM that may provide an understanding of the T1DM pathogenesis and assist in identifying potential therapeutic targets. Although GWAS

have provided new insights for assessing the association between SNPs and disease statistically, limitations still emerge. One of the limitations is that GWAS only account for a modest proportion of the missing heritability, including T1DM, which has complex traits. Additionally, current genome-wide complex trait analyses indicate that SNPs may account for one-third to two-thirds of the heritability of the majority of complex traits. One possible explanation is due the SNPs fall short of the strict significance level; thus, these SNPs with a minor effect may be ignored [72]. Furthermore, the clinical predictive value of GWAS is limited. Thus, screening population for the genetics is not practicable due to the highly exceed number of false positives compared to true positives [72, 73].

Although the current study did not cross-validate with any additional databases other than those stated, it was assessed using three large datasets, namely GWAS, SNPnexus, and GTEx Portal. Thus, this study can provide information about the missense genetic variations related to T1DM disease, which may aid in the identification of novel treatment targets. However, considering the limitations, further preclinical and clinical studies with larger sample sizes and adequate study designs are necessary to validate these findings.

These further studies may be helpful for analysing more complex interactions with phenotypes and complex diseases, as well as translating these discoveries into clinical practice. This includes analysing a druggable target that is related to the pathophysiology and genetics of T1DM. Following that, identifying and comprehending the disease's heterogeneity can contribute to tailored therapy that improves treatment success, resulting in a better long-term outcome for individuals with T1DM [74].

## 5 Summary

This present bioinformatics study displayed that the missense genetic variants of *TYK2*, *IFIH1*, and *COL4A3* genes were associated with T1DM disease. The loss-of-function of *TYK2* rs34536443 and decreased level of *IFIH1* are known for the protective function of T1DM. Similarly, the *COL4A3* rs55703767 is significantly related to protection from diabetic nephropathy. Therefore, these genes are excellent candidates for further investigation that may provide new insights into the pathogenesis and therapeutic target in T1DM.

### Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Authors' contributions

Conceptualization: RDY, DNP. Data curation: DNP. Formal analysis: RDY, DNP. Funding acquisition: –Methodology: RDY, LMI. Writing—original draft: RDY, DNP. Writing—review and editing: LMI.

## Declarations

### Ethics approval and consent to participate

Ethical review is not applicable for this study. This study analysed datasets from open sources.

### Competing interests

No potential competing interests relevant to this article were reported.

Received: 29 November 2023 Accepted: 3 March 2024

Published online: 26 June 2024

## References

- Katsarou A, Gudbjörnsdóttir S, Rawshani A, Dabelea D, Bonifacio E, Anderson BJ, et al. Type 1 diabetes mellitus. *Nat Rev Dis Primers*. 2017;3(1):17016.
- Yudhani RD, Nugrahaningsih DAA, Sholikah EN, Mustofa M. The molecular mechanisms of hypoglycemic properties and safety profiles of *Swietenia macrophylla* seeds extract: a review. *Open Access Maced J Med Sci*. 2021;9(F):370–88.
- Diaz-Valencia PA, Bougnères P, Valleron A-J. Global epidemiology of type 1 diabetes in young adults and adults: a systematic review. *BMC Public Health*. 2015;15(1):255.
- Guo SJ, Shao H. Growing global burden of type 1 diabetes needs multitiered precision public health interventions. *Lancet Diabetes Endocrinol*. 2022;10(10):688–9.
- Lind M, Svensson A-M, Kosiborod M, Gudbjörnsdóttir S, Pivodic A, Wedel H, et al. Glycemic control and excess mortality in type 1 diabetes. *N Engl J Med*. 2014;371(21):1972–82.
- von Scholten BJ, Kreiner FF, Gough SCL, von Herrath M. Current and future therapies for type 1 diabetes. *Diabetologia*. 2021;64(5):1037–48.
- Yahaya T, Salisu T. Genes predisposing to type 1 diabetes mellitus and pathophysiology: a narrative review. *arXiv preprint arXiv:210106680*. 2021.
- Klak M, Gomółka M, Kowalska P, Cichoń J, Ambrożkiewicz F, Serwańska-Swiętek M, et al. Type 1 diabetes: genes associated with disease development. *Cent Eur J Immunol*. 2020;45(4):439–53.
- Krischer JP, Lynch KF, Schatz DA, Ilonen J, Lernmark Å, Hagopian WA, et al. The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study. *Diabetologia*. 2015;58(5):980–7.
- Nyaga DM, Vickers MH, Jefferies C, Perry JK, O'Sullivan JM. Type 1 diabetes mellitus-associated genetic variants contribute to overlapping immune regulatory networks. *Front Genet*. 2018;9:535.
- Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Mangano C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47(D1):D1005–12.
- Cheng J, Novati G, Pan J, Bycroft C, Żemgulytė A, Applebaum T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381(6664):eadg7492.
- Ranganathan P, Aggarwal R, Pramesh CS. Common pitfalls in statistical analysis: odds versus risk. *Perspect Clin Res*. 2015;6(4):222–4.
- Irham LM, Adikusuma W, Lolita L, Puspitaningrum AN, Afief AR, Sarasmita MA, et al. Investigation of susceptibility genes for chickenpox disease across multiple continents. *Biochem Biophys Res*. 2023;33:101419.
- Satria RD, Irham LM, Adikusuma W, Puspitaningrum AN, Afief AR, El Khair R, et al. Identification of druggable genes for multiple myeloma based on genomic information. *Genomics Inform*. 2023;21(3):e31.
- Yudhani RD, Pakha DN, Suyatni S, Irham LM. Identifying pathogenic variants related to systemic lupus erythematosus by integrating genomic databases and a bioinformatic approach. *Genomics Inform*. 2023;21(3):e37.
- Di Leo G, Sardanelli F. Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. *Eur Radiol Exp*. 2020;4(1):18.
- Oscanoa J, Sivapalan L, Gadaleta E, Dayem Ullah AZ, Lemoine Nicholas R, Chelala C. SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). *Nucleic Acids Res*. 2020;48(W1):W185–92.

19. Dayem Ullah AZ, Oscanoa J, Wang J, Nagano A, Lemoine NR, Chelala C. SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res.* 2018;46(W1):W109–13.
20. Dayem Ullah AZ, Lemoine NR, Chelala C. A practical guide for the functional annotation of genetic variations using SNPnexus. *Brief Bioinform.* 2013;14(4):437–47.
21. Dayem Ullah AZ, Lemoine NR, Chelala C. SNPnexus: a web server for functional annotation of novel and publicly known genetic variants (2012 update). *Nucleic Acids Res.* 2012;40(W1):W65–70.
22. Chelala C, Khan A, Lemoine NR. SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics.* 2008;25(5):655–61.
23. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248–9.
24. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;Chapter 7:Unit7.20.
25. Belsare S, Levy-Sakin M, Mostovoy Y, Durinck S, Chaudhuri S, Xiao M, et al. Evaluating the quality of the 1000 genomes project data. *BMC Genomics.* 2019;20(1):620.
26. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68–74.
27. Consortium G. The genotype-tissue expression (GTEx) project. *Nat Genet.* 2013;45(6):580–5.
28. Puspitaningrum AN, Perwitasari DA, Adikusuma W, D'jalilah GN, Dania H, Maliza R, et al. Integration of genomic databases and bioinformatic approach to identify genomic variants for Sjogren's syndrome on multiple continents. *J Pharm Sci.* 2022;19(2):71–81.
29. Marroqui L, Dos Santos RS, Fløyel T, Grieco FA, Santin I, Op de beeck A, et al. TYK2, a candidate gene for type 1 diabetes, modulates apoptosis and the innate immune response in human pancreatic  $\beta$ -cells. *Diabetes.* 2015;64(11):3808–17.
30. Shapiro MR, Thirawatanaonond P, Peters L, Sharp RC, Ogundare S, Posgai AL, et al. De-coding genetic risk variants in type 1 diabetes. *Immunol Cell Biol.* 2021;99(5):496–508.
31. Kasper S, Kindler T, Sonnenschein S, Breitenbuecher F, Böhmer FD, Huber C, et al. Cross-inhibition of interferon-induced signals by GM-CSF through a block in Stat1 activation. *J Interferon Cytokine Res.* 2007;27(11):947–60.
32. Morimoto J, Yoneyama H, Shimada A, Shigihara T, Yamada S, Oikawa Y, et al. CXC chemokine ligand 10 neutralization suppresses the occurrence of diabetes in nonobese diabetic mice through enhanced beta cell proliferation without affecting insulinitis. *J Immunol.* 2004;173(11):7017–24.
33. Chandra V, Ibrahim H, Halliez C, Prasad RB, Vecchio F, Dwivedi OP, et al. The type 1 diabetes gene TYK2 regulates  $\beta$ -cell development and its responses to interferon- $\alpha$ . *Nat Commun.* 2022;13(1):6363.
34. Dendrou CA, Cortes A, Shipman L, Evans HG, Attfield KE, Jostins L, et al. Resolving TYK2 locus genotype-to-phenotype differences in autoimmunity. *Sci Transl Med.* 2016;8(363):363ra149.
35. Yang H, Wang Z, Xu K, Gu R, Chen H, Yu D, et al. IFIH1 gene polymorphisms in type 1 diabetes: genetic association analysis and genotype-phenotype correlation in Chinese Han population. *Autoimmunity.* 2012;45(3):226–32.
36. Downes K, Pekalski M, Angus KL, Hardy M, Nutland S, Smyth DJ, et al. Reduced expression of IFIH1 is protective for type 1 diabetes. *PLoS One.* 2010;5(9):e12646.
37. Bouças AP, Oliveira Fdos S, Canani LH, Crispim D. The role of interferon induced with helicase C domain 1 (IFIH1) in the development of type 1 diabetes mellitus. *Arq Bras Endocrinol Metabol.* 2013;57(9):667–76.
38. Kawashima H, Ihara T, Ioi H, Oana S, Sato S, Kato N, et al. Enterovirus-related type 1 diabetes mellitus and antibodies to glutamic acid decarboxylase in Japan. *J Infect.* 2004;49(2):147–51.
39. Moya-Suri V, Schlosser M, Zimmermann K, Rjasanowski I, Gürtler L, Mentel R. Enterovirus RNA sequences in sera of schoolchildren in the general population and their association with type 1-diabetes-associated autoantibodies. *J Med Microbiol.* 2005;54(Pt 9):879–83.
40. Shibasaki S, Imagawa A, Tauriainen S, Iino M, Oikarinen M, Abiru H, et al. Expression of toll-like receptors in the pancreas of recent-onset fulminant type 1 diabetes. *Endocr J.* 2010;57(3):211–9.
41. Tanaka S, Nishida Y, Aida K, Maruyama T, Shimada A, Suzuki M, et al. Enterovirus infection, CXC chemokine ligand 10 (CXCL10), and CXCR3 circuit: a mechanism of accelerated beta-cell failure in fulminant type 1 diabetes. *Diabetes.* 2009;58(10):2285–91.
42. Smyth DJ, Cooper JD, Bailey R, Field S, Burren O, Smink LJ, et al. A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nat Genet.* 2006;38(6):617–9.
43. Jermendy A, Szatmári I, Laine AP, Lukács K, Horváth KH, Körner A, et al. The interferon-induced helicase IFIH1 Ala946Thr polymorphism is associated with type 1 diabetes in both the high-incidence Finnish and the medium-incidence Hungarian populations. *Diabetologia.* 2010;53(1):98–102.
44. Chistiakov DA, Voronova NV, Savost'Anov KV, Turakulov RI. Loss-of-function mutations E6 27X and I923V of IFIH1 are associated with lower poly(I:C)-induced interferon- $\beta$  production in peripheral blood mononuclear cells of type 1 diabetes patients. *Hum Immunol.* 2010;71(11):1128–34.
45. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science.* 2009;324(5925):387–9.
46. Salem RM, Todd JN, Sandholm N, Cole JB, Chen WM, Andrews D, et al. Genome-wide association study of diabetic kidney disease highlights biology involved in glomerular basement membrane collagen. *J Am Soc Nephrol.* 2019;30(10):2000–16.
47. Xia L, Cao Y, Guo Y, Ba G, Luo Q, Shi H, et al. A novel heterozygous mutation of the COL4A3 gene causes a peculiar phenotype without hematuria and renal function impairment in a Chinese family. *Dis Markers.* 2019;2019:8705989.
48. Heidet L, Arrondel C, Forestier L, Cohen-Solal L, Mollet G, Gutierrez B, et al. Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome. *J Am Soc Nephrol.* 2001;12(1):97–106.
49. Bereda G. Risk factors and pathogenesis of diabetic nephropathy. *Int J Diabetes Metab Disord.* 2022;7(1):05, 7.
50. Gross JL, de Azevedo MJ, Silveiro SP, Canani LH, Caramori ML, Zelmanovitz T. Diabetic nephropathy: diagnosis, prevention, and treatment. *Diabetes Care.* 2005;28(1):164–76.
51. Kochenborger APB, Ripoll APK, de Souza BM, Nique PS, Sortica DA, Cardoso N, et al. Hba1c levels are increased in patients with gestational diabetes carrying the T/T genotype of the rs1990760 polymorphism in the IFIH1 gene. *Diabetol Metab Syndr.* 2015;7(1):A75.
52. Nagafuchi S, Kamada-Hibio Y, Hirakawa K, Tsutsu N, Minami M, Okada A, et al. TYK2 promoter variant and diabetes mellitus in the Japanese. *EBioMedicine.* 2015;2(7):744–9.
53. Mori H, Takahashi H, Mine K, Higashimoto K, Inoue K, Kojima M, et al. TYK2 promoter variant is associated with impaired insulin secretion and lower insulin resistance in Japanese type 2 diabetes patients. *Genes (Basel).* 2021;12(3):400.
54. Saravani S, Yari D, Saravani R, Azadi AC. Association of COL4A3 (rs55703767), MMP-9 (rs17576) and TIMP-1 (rs6609533) gene polymorphisms with susceptibility to type 2 diabetes. *Biomed Rep.* 2017;6(3):329–34.
55. Majeed SM, Omara AM, Al-Koofee DAF. Association of interferon-induced helicase (IFIH1) gene polymorphism rs1990760 with type two diabetes mellitus in Iraqi population. *Meta Gene.* 2021;30:100952.
56. Mine K, Yoshikai Y, Takahashi H, Mori H, Anzai K, Nagafuchi S. Genetic susceptibility of the host in virus-induced diabetes. *Microorganisms.* 2020;8(8):1133.
57. Ortega-Contreras B, Armella A, Appel J, Mennicken D, Araya J, González M, et al. Pathophysiological role of genetic factors associated with gestational diabetes mellitus. *Front Physiol.* 2022;13:769924.
58. Karvonen M, Viik-Kajander M, Moltchanova E, Libman I, LaPorte R, Tuomilehto J. Incidence of childhood type 1 diabetes worldwide. *Diabetes Mondiale (DiaMond) Project Group. Diabetes Care.* 2000;23(10):1516–26.
59. Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of type 1 diabetes. *Endocrinol Metab Clin North Am.* 2010;39(3):481–97.

60. Dabelea D, Bell RA, D'Agostino RB Jr, Imperatore G, Johansen JM, Linder B, et al. Incidence of diabetes in youth in the United States. *JAMA*. 2007;297(24):2716–24.
61. Mobasser M, Shirmohammadi M, Amiri T, Vahed N, Hosseini Fard H, Ghajzadeh M. Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. *Health Promot Perspect*. 2020;10(2):98–115.
62. Rich SS. Genetic contribution to risk for diabetic kidney disease. *Clin J Am Soc Nephrol*. 2018;13(8):1135–7.
63. Smyth LJ, Kilner J, Nair V, Liu H, Brennan E, Kerr K, et al. Assessment of differentially methylated loci in individuals with end-stage kidney disease attributed to diabetic kidney disease: an exploratory study. *Clin Epigenetics*. 2021;13(1):99.
64. Basu M, Wang K, Ruppin E, Hannehalli S. Predicting tissue-specific gene expression from whole blood transcriptome. *Sci Adv*. 2021;7(14):eabd6991.
65. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. *Nat Genet*. 2013;45(6):580–5.
66. GeneCards. The human gene database. TYK2 gene - tyrosine kinase 2. Available from: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=TYK2#expression>. Cited 2023 23 June.
67. Poelzl A, Lassnig C, Tangermann S, Hromadová D, Reichart U, Gawish R, et al. TYK2 licenses non-canonical inflammasome activation during endotoxemia. *Cell Death Differ*. 2021;28(2):748–63.
68. UniProt. IFIH1\_Mouse. Available from: <https://www.uniprot.org/uniprot/otkb/Q8R5F7/entry>. Cited 2023 23 June.
69. Bronte V, Pittet MJ. The spleen in local and systemic regulation of immunity. *Immunity*. 2013;39(5):806–18.
70. NCBI. COL4A3 collagen type IV alpha 3 chain [Homo sapiens (human)]. Available from: <https://www.ncbi.nlm.nih.gov/gene/1285>. Cited 2023 1 July.
71. Mohandes S, Doke T, Hu H, Mukhi D, Dhillon P, Susztak K. Molecular pathways that drive diabetic kidney disease. *J Clin Invest*. 2023;133(4):e165654.
72. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467–84.
73. Manolio TA. Bringing genome-wide association findings into clinical use. *Nat Rev Genet*. 2013;14(8):549–58.
74. Akil AA-S, Yassin E, Al-Maraghi A, Aliyev E, Al-Malki K, Fakhro KA. Diagnosis and treatment of type 1 diabetes at the dawn of the personalized medicine era. *J Transl Med*. 2021;19(1):137.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.