

https://dspace.org/

The association of HLA-G with type 1 diabetes in the South African black population

Dzimbanhete, Charmaine Ruvimbo

2024
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/48051
Dzimbanhete, Charmaine Ruvimbo . (2024). The association of HLA-G with type 1 diabetes in the South African black population [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/48051

Downloaded from a DSpace repository. https://dspace.org/

The association of HLA-G with type 1 diabetes in the South African black population

Charmaine Ruvimbo Dzimbanhete



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science in Medicine

Johannesburg, 2024

Declaration

I, Charmaine Ruvimbo Dzimbanhete, declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in Medicine by Research (Chemical Pathology) at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



Signature

On the _21st___day of ___November_____2024

Dedication

I would like to dedicate this dissertation to:

El Roi, muri Mwari anogona, Mwari akatendeka, Mwari wenyasha

Mubatseri wangu, Mweya Mutsvene

My parents Everesto and Betty Dzimbanhete, for their endless love and support

Presentations arising from this study

Poster presentations:

1. *55th Society of Endocrinology, Metabolism and Diabetes of South Africa Congress (Cape Town)*

   Dzimbanhete CR, Cave EM, Bhola S, Crowther NJ, Padoa CJ (2022) The HLA-G 14 bp indel polymorphism is associated with type 1 diabetes in a South African black population.
2. *WITS University Faculty of Health Sciences Research Day & Postgraduate Expo (Johannesburg)*

   Dzimbanhete CR, Cave EM, Bhola S, Crowther NJ, Padoa CJ (2022) The HLA-G 14 bp indel polymorphism is associated with type 1 diabetes in a South African black population.
3. *WITS University Faculty of Health Sciences Research Day & Postgraduate Expo (Johannesburg)*

   Dzimbanhete CR, Cave EM, Bhola S, Crowther NJ, Padoa CJ (2024) The association of *HLA-G* with type 1 diabetes in a South African black population.

Acknowledgements

I would like to acknowledge and thank my supervisors, Dr Carolyn Padoa and Dr Eleanor Cave, for their guidance and continuous and unwavering support throughout my project.

Thank you to the National Health Laboratory Services Research Trust for funding my project.

Thank you to all the participants who volunteered to take part in the study and to Sureka Bhola, who reached out to all the prospective participants, thank you for all the hard work you have put into recruiting the participants.

A special thanks to my family and friends for their love and encouragement.

Abstract

Background: Type 1 diabetes mellitus (T1D) is an autoimmune disease that is characterized by a T cell mediated destruction of pancreatic β-cells which leads to an absolute insulin deficiency. The cause of T1D is multifactorial and *HLA-G*, a non-classical *HLA* class Ib gene, was identified as an independent locus for disease susceptibility. *HLA-G* has limited allelic and protein variability, and unique alternative splicing patterns that produce multiple isoforms. It has immune-suppressive properties which include altering cytokine secretion, downregulating the production of $CD4^+$ T cells, inhibiting antigen-presenting cells, and inhibiting the activity and inducing apoptosis of cytotoxic $CD8^+$ T cells and natural killer (NK) cells. The HLA-G molecule is constitutively expressed in the pancreas and is upregulated on the surface of islet cells that have been stimulated to secrete insulin. When there are more HLA-G molecules on the cell-surface, $CD8^+$ T cell activation is suppressed, creating immune tolerance within the pancreas. Decreased expression of *HLA-G* therefore results in decreased tolerance to pancreatic autoantigens, and potentially the initiation and progression of T1D. Polymorphisms in the *HLA-G* gene, affect isoform production, peptide binding, mRNA expression and consequently the immunoregulatory functions of HLA-G. Polymorphisms in the 3' untranslated region (UTR) of the *HLA-G* gene (a 14 bp insertion/deletion polymorphism and rs9380142) have been associated with an increased risk of developing T1D and a younger age at diagnosis of the disease. Therefore, the aim of this study was to sequence the *HLA-G* gene in black South Africans with T1D (cases) and controls and to determine whether polymorphisms in the *HLA-G* gene are associated with age at diagnosis and/or T1D in the South African black population. In addition, this study aimed to determine the effect of *HLA-G* genotypes on serum HLA-G concentrations.

Methods: A total of 263 cases were previously recruited from the greater Johannesburg area. The eight exons of the *HLA-G* gene were sequenced and compared in 20 cases and 20 controls. The *HLA-G* sequences were aligned with a reference sequence (GRCh38.p14; NC_000006.12) to identify variants that had significantly different frequencies between cases and controls. Two variants in exon 8, rs1063320 and rs1710, were selected for further analysis. Participants were genotyped for the *HLA-G* rs1063320 and rs1710 polymorphisms using PCR-RFLP. In

addition, participants were genotyped for the 14 bp indel using conventional PCR and the rs9380142 polymorphism using a predesigned Applied Biosystems TaqMan® SNP Genotyping assay. The concentration of serum HLA-G was measured by an enzyme linked immunosorbent assay. Statistical analysis was performed using Statistica software.

Results: In our cohort the mean age at diagnosis of T1D was 19.3 ± 10.1 years with a median duration of 6.00 [2.00; 12.0] years. Blood glucose concentrations ($p<0.001$), systolic blood pressure (SBP) ($p=0.021$) and serum HLA-G concentrations ($p<0.001$) were significantly higher in cases than controls. There was no significant difference in the genotypic and allelic frequencies between cases and controls for any of the *HLA-G* polymorphisms studied ($p>0.05$). Similarly, there was no association between *HLA-G* polymorphisms and HLA-G concentrations ($p>0.05$). In a univariate analysis none of the four polymorphisms studied were associated with age at diagnosis, however, on multivariate analysis the 14 bp insertion/insertion genotype was associated with an earlier age at diagnosis (0.90 (0.81-1.00); $p=0.042$). Participants with T1D were 5.2 times more likely to have higher HLA-G concentrations compared to controls ($p<0.001$). In addition, participants with the 14 bp deletion/deletion genotype had a SBP, 4.00 mmHg higher than those with the insertion allele ($p=0.036$). Furthermore, for every one-year increase in age and one unit increase in BMI, the SBP increased by 0.412 mmHg ($p<0.001$) and 0.67 mmHg ($p<0.001$), respectively.

Conclusion: The rs1063320, rs1710, 14 bp indel and rs9380142 polymorphisms are not associated with T1D in the black South African population. The association of the 14 bp homozygous insertion with an earlier age at diagnosis supports the involvement of *HLA-G* in the pathogenesis of T1D. Serum HLA-G concentrations were higher in cases than controls. Our findings contrast with that reported in literature where the 14 bp deletion and lower HLA-G concentrations have been associated with age at diagnosis and T1D, respectively. It has been hypothesised that increased HLA-G levels cause apoptosis of $CD8^+$ cytotoxic T cells and natural killer cells, thus protecting pancreatic β-cells from autoimmune destruction. However, our study does not support this hypothesis and future studies should be performed to ascertain the involvement of *HLA-G* in T1D pathogenesis in the black South African population.

Table of contents

Declaration i

Dedication ii

Presentations arising from this study iii

Acknowledgements iv

Abstract v

Table of contents vii

List of abbreviations, symbols, and units xi

List of figures xiv

List of tables xvi

1. Introduction 1

1.1 Diabetes mellitus 1

1.1.1 Diagnosis of diabetes 1

1.1.2 Classification of diabetes 1

1.2 Type 1 diabetes mellitus 2

1.2.1 Pathogenesis of T1D 2

1.2.2 Epidemiology of T1D 4

1.3 Genetic susceptibility to T1D 4

1.3.1 The human leukocyte antigen complex and its role in T1D 4

1.3.2 The human leukocyte antigen-G 5

1.3.3 The *HLA-G* gene and protein structure 5

1.3.4 The functions of HLA-G 7

1.3.5 The role of HLA-G in T1D 7

1.4 *HLA-G* polymorphisms and T1D 9

1.4.1 The HLA-G 14 bp insertion/deletion polymorphism 9

1.4.2 The *HLA*-G rs9380142 (+3187 A>G) polymorphism 11

1.5 Aim and study objectives 12
2. Methods 13
2.1 Study participants 13
2.1.1 Sample size 13
2.2 Sequencing of the *HLA-G* gene 14
2.2.1 PCR amplification 14
2.2.2 Agarose gel electrophoresis 15
2.2.3 DNA purification for exon 4-5 and exon 6-8 15
2.2.4 Sequencing of PCR amplicons 16
2.3 Genotyping participants for the *HLA-G* rs1063320 and rs1710 polymorphisms 17
2.3.1 PCR amplification 17
2.3.2 Agarose gel electrophoresis 17
2.3.3 Restriction digest 18
2.4 Genotyping participants for the *HLA-G* 14 bp I/D polymorphism 19
2.4.1 PCR amplification 19
2.4.2 Agarose gel electrophoresis 20
2.5 Genotyping participants for the *HLA-G* rs9380142 polymorphism 20
2.5.1 Real-time PCR amplification 20
2.6 Confirmation of genotypes by sequencing 21
2.7 Measurement of serum HLA-G concentrations 21
2.7.1 Preparation of reagents and participant serum samples 21
2.7.2 Preparation of HLA-G standards 22
2.7.3 Assay procedure 22
2.8 Statistical analysis 23
3. Results 25
3.1 Anthropometric and clinical characteristics of the study population 25

3.2 Sequencing of *HLA-G* exons 26

3.3 Determining the *HLA-G* gene genotypes for rs1063320, rs1710, 14 bp indel and rs9380142 28

3.3.1 Detection of the *HLA*-G rs1063320and rs1710 SNPs 28

3.3.2 Detection of the *HLA-G* 14 bp insertion/deletion polymorphism 30

3.3.3 Detection of the *HLA-G* rs9380142 SNP 31

3.3.4 Sequencing 33

3.4 Genotypic and allelic frequencies of the *HLA-G* gene polymorphisms 35

3.5 The association of *HLA-G* gene polymorphisms with clinicopathological variables in the whole population 37

3.6 The association of *HLA-G* gene polymorphisms with clinicopathological variables in participants with T1D 39

3.7 The association of *HLA-G* gene polymorphisms in participants with T1D with an age at diagnosis before and after 21 years of age 43

3.8 Logistic regression model determining variables associated with T1D 45

3.9 Logistic regression model determining variables associated with age at diagnosis category 45

3.10 Multivariable linear regression model to determine factors affecting serum HLA-G concentrations 46

4. Discussion 47

4.1 Participant characteristics 47

4.2 Identification of variants in the *HLA-G* gene of black South African participants 48

4.3 Allelic and genotypic frequencies of the four *HLA-G* polymorphisms 49

4.4 The *HLA-G* polymorphisms are not associated with T1D in the black South African population 49

4.5 Serum HLA-G levels are associated with T1D in the black South African population 51

4.6 The 14 bp insertion is associated with an earlier age at diagnosis of T1D in the black South African population 52
4.7 The 14 bp indel polymorphism is associated with SBP in the South African black population. 53
4.8 Age, BMI and T1D status are associated with SBP in the South African black population 54
4.9 Limitations of the study 55
5. Conclusions 56
6. References 57
7. Appendices 65
7.1 Appendix A: Stock solutions and dilution preparations for PCR 65
7.2 Appendix B: Ethics certificates 66
7.3 Appendix C: *HLA-G* variants identified from sequencing analysis 68
7.4 Appendix D: ELISA standard curve 87
7.5 Appendix E: Hardy-Weinberg equilibrium 88
7.6 Appendix F: PCR gel images of amplified exon regions 90
7.7 Appendix G: Turn-it-in report 91

List of abbreviations, symbols, and units

A	Adenine
Ab	Antibody
ADB	Agarose dissolving buffer
Ag	Antigen
ANOVA	Analysis of variance
APC	Antigen-presenting cell
AU	Adenine and uridine
BMI	Body mass index
bp	Base pair
C	Cytosine
CD	Cluster of differentiation
conc	concentration
D	Deletion
DBP	Diastolic blood pressure
DC	Dendritic cell
DC-10	Interleukin-10 secreting dendritic cells
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
e.g.	For example
ELISA	Enzyme-linked immunosorbent assay
G	Guanine
GAD65	65 Kilodalton isoform of Glutamic acid decarboxylase
GWAS	Genome wide association studies
HbA1c	Glycated haemoglobin
HLA	Human leukocyte antigen
HNP	Histologically normal pancreatic
HRP	Horseradish peroxidase
HWE	Hardy-Weinberg equilibrium
I	Insertion
i.e.	That is
IDF	International Diabetes Federation

IFN-γ	Interferon gamma
IL	Interleukin
ILT2	Immunoglobulin-like transcription type 2
ILT4	Immunoglobulin-like transcription type 4
Indel	Insertion/deletion
KIR2DL4	Killer immunoglobulin-like receptor 2DL4
Mϕ	Macrophages
MHC	Major histocompatibility complex
MHCG	Major histocompatibility complex class I G
mRNA	Messenger ribonucleic acid
miRNA	microRNA
NCBI	National Center for Biotechnology Information
NK	Natural killer
NTC	Non-template control
OD	Optical density
OGTT	Oral glucose tolerance test
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RSS	Reference Standard and Sample
SANBS	South African National Blood Services
SBP	Systolic blood pressure
SEMDSA	Society of Endocrinology, Metabolism and Diabetes of South Africa
SD	Standard deviation
sHLA-G	Soluble human leukocyte antigen-G
SNP	Single nucleotide polymorphism
T1D	Type 1 diabetes
T2D	Type 2 diabetes
TBE	Tris–Borate-EDTA
TE	Tris and EDTA
TGF-β	Transforming growth factor-beta
Th1	T helper 1

Th2	T helper 2
USA	United States of America
UTR	Untranslated region
UV	Ultraviolet
V	Volts
VEGF	vascular endothelial growth factor
vs	Versus
α	Alpha
β	Beta
°C	Degrees Celsius
Chi^2 (χ^2)	Chi square
dH_2O	Distilled water
diH_2O	Deionized water
g	Relative centrifugal force
kg/m^2	Kilogram per metre squared
L	Litre
ml	Millilitre
mg	Milligram
mmHg	Millimetre of mercury
mmol	Millimole
n	Number
ng	Nanogram
µl	Microlitre
µM	Micromolar
U/ml	Unit per millilitre
±	Plus/minus

List of figures

Figure 1.1: A schematic representation of the proposed pathogenesis of T1D (Wallberg and Cooke, 2013). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended). 3
Figure 1.2: A schematic diagram of the position of the HLA region on the short arm of chromosome 6 including the classical and non-classical genes (Arnaiz-Villena et al., 2021). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended). 5
Figure 1.3: Schematic representation of the protein products coded by the HLA-G gene (Krijgsman et al., 2020). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended). 6
Figure 1.4: Schematic representation of the 14 bp HLA-G indel and its effect on HLA-G expression (Martin-Villa et al., 2022). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended). 10
Figure 1.5: The effect of the HLA-G +3187 (A>G) SNP on the expression of the HLA-G gene (adapted from Martin-Villa et al., 2022). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended). 11
Figure 2.1: Dilution series of HLA-G standards 22
Figure 3.1: A 1 % agarose gel showing the PCR products obtained from the amplification of a 255 bp region of the HLA-G gene containing the rs1063320 and rs1710 polymorphisms 28
Figure 3.2: A 2% agarose gel showing the restriction fragment profile of the HLA-G rs1063320 polymorphism PCR products digested with ApaLI 29
Figure 3.3: A 2% agarose gel showing the restriction profile of the HLA-G rs1710 polymorphism PCR products digested with PstI 29
Figure 3.4: A 3% agarose gel image of PCR products obtained from the amplification of the HLA-G gene 3' UTR containing the 14 bp indel 30
Figure 3.5: Amplification plot for the *HLA-G* rs9380142 SNP. 31
Figure 3.6: Allelic discrimination plot for the HLA-G rs9380142 SNP 32
Figure 3.7: Chromatograms obtained from the amplification of the HLA-G gene containing the rs1063320 polymorphism. This SNP was sequenced in the forward and reverse direction. 33

Figure 3.8: Chromatograms obtained from the amplification of the HLA-G gene containing the rs1710 polymorphism. This SNP was sequenced in the forward and reverse direction. 34
Figure 3.9: Chromatograms obtained from the amplification of the *HLA-G* gene containing the 14 bp indel polymorphism 35
Figure 7.1: The four-parameter logistic (standard) curve used to calculate serum concentrations of HLA-G in different samples using OD determined from the ELISA assay 87
Figure 7.2: PCR products from amplification of all HLA-G exons run on a 1% agarose gel 90

List of tables

Table 1.1: Aetiological classification of diabetes (Adapted from American diabetes association, 2023) 2
Table 2.1: Primer sequences flanking exons of the HLA-G gene and the corresponding PCR amplicon fragment sizes 14
Table 2.2: Primer sequences flanking the region containing the HLA-G rs1063320 and rs1710 polymorphisms 17
Table 2.3: Reagent volumes for the digestion of rs1063320 and rs1710 PCR products 18
Table 2.4: PCR-RFLP fragment sizes and their corresponding genotypes for the HLA-G rs1063320 and rs1710 polymorphisms..... 19
Table 2.5: Primer sequences flanking the region containing the 14 bp indel 19
Table 2.6: PCR amplicon fragment sizes and their corresponding genotypes for the 14 bp indel..... 20
Table 3.1: Comparison of anthropometric and clinical characteristics between the cases and controls groups 26
Table 3.2: HLA-G variants and their frequencies in cases and controls 27
Table 3.3: Genotypic and allelic frequencies of the HLA-G gene polymorphisms in participants with T1D and controls 36
Table 3.4: Association of HLA-G rs1063320 genotypes with clinicopathological variables..... 37
Table 3.5: Association of HLA-G rs1710 genotypes with clinicopathological variables 38
Table 3.6: Association of HLA-G 14 bp indel genotypes with clinicopathological variables..... 38
Table 3.7: Association of HLA-G rs9380142 genotypes with clinicopathological variables..... 39
Table 3.8: Association of the HLA-G rs1063320 genotypes with clinicopathological variables in participants with T1D..... 40
Table 3.9: Association of the HLA-G rs1710 genotypes with clinicopathological variables in participants with T1D..... 41

Table 3.10: Association of the HLA-G 14 bp indel genotypes with clinicopathological variables in participants with T1D 42
Table 3.11: Association of the HLA-G rs9380142 genotypes with clinicopathological variables in participants with T1D 43
Table 3.12: Comparison of genotypic and allelic frequencies of HLA-G polymorphisms between participants with T1D and an early and late age at diagnosis 44
Table 3.13: Logistic regression analysis for the determinants of T1D status 45
Table 3.14: Logistic regression analysis for the determinants of age at diagnosis of T1D 46
Table 3.15: Multivariable linear regression analysis for determinants of systolic blood pressure 46
Table 7.1: Hardy-Weinberg equilibrium 88
Table 7.2: Comparison of HLA-G polymorphism frequencies between our cohort and the 1000 genome African population 89

1. Introduction

1.1 Diabetes mellitus

Diabetes mellitus is a chronic metabolic condition that manifests as a consequence of the body's inability to produce or effectively use insulin to lower blood glucose levels (Wang *et al*., 2018). Insulin is a peptide hormone that modulates the metabolism of carbohydrates, proteins, and lipids by promoting the uptake of glucose in the liver, skeletal muscles, and adipose tissue (Galicia-Garcia *et al.,* 2020). The mechanisms involved in glucose homeostasis are tightly regulated as any defect in the feedback loops can lead to hyperglycaemia and subsequently, diabetes (Mashal *et al*., 2021).

1.1.1 Diagnosis of diabetes

An individual is diagnosed with diabetes when they meet one of the following criteria: fasting plasma glucose ≥ 7 mmol/L, two-hour plasma glucose ≥11.1 mmol/L, HbA1c ≥ 6.5 % or a random plasma glucose concentration >11.1 mmol/L in the presence of symptoms of hyperglycaemia, polyuria (excessive urine production), sudden weight loss, polydipsia (excessive thirst), polyphagia (extreme hunger), blurred vision, lack of energy, and diabetic ketoacidosis (SEMDSA type 2 diabetes guidelines expert committee, 2017).

1.1.2 Classification of diabetes

Based on the aetiology and pathogenesis of the disease, diabetes is divided into four main types: type 1 diabetes (T1D), type 2 diabetes (T2D), gestational diabetes and other specific types of diabetes which include monogenic diabetes (Table 1.1) (Solis-Herrera *et al*., 2016).

Table 1.1: Aetiological classification of diabetes (Adapted from American diabetes association, 2023)

Types of diabetes	Aetiology
Type 1 diabetes	Immune mediated destruction of β-cells resulting in absolute insulin deficiency
Type 2 diabetes	Insulin resistance with insulin secretory defects
Gestational diabetes	Glucose intolerance with first onset during pregnancy
Other specific types	Due to a range of causes including genetic defects of β-cell function and insulin action, diseases of the exocrine pancreas, endocrinopathies, drug or chemical induced, infections, and other genetic syndromes

1.2 Type 1 diabetes mellitus

Type 1 diabetes mellitus (T1D) is an autoimmune disease that is characterized by a T cell mediated destruction of pancreatic β-cells that leads to an absolute insulin deficiency (Amodio *et al*., 2021). The cause of T1D is multifactorial (Silva *et al*., 2016), however, the exact mechanism involved in the initiation and progression of the β-cell destruction is largely unknown.

1.2.1 Pathogenesis of T1D

It has been proposed that when a genetically susceptible individual is exposed to an environmental trigger such as a viral infection, especially enteroviruses, T cells and macrophages infiltrate the pancreatic islets of Langerhans (Figure 1.1). The presence of these mononuclear cells promotes the initial destruction of β-cells, which leads to the release of autoantigens (e.g. insulin and the 65 kilodalton isoform of glutamic acid decarboxylase (GAD65)) that trigger the following autoimmune response (Kukreja and Maclaren, 1999; Gerasimou *et al*., 2016). The autoantigens are presented to naïve $CD4^+$ T helper (Th) cells by macrophages (Mϕ) or dendritic cells (DC) via the major histocompatibility complex (MHC) class II molecules (Kongsbak *et al*., 2013). This interaction coupled with the release of interleukin (IL)-12 by the antigen-presenting cells (APCs), stimulates the differentiation of naïve $CD4^+$ Th cells into activated Th 1 cells which secrete IL-2 and interferon gamma (IFN-ϒ). IFN-ϒ then activates other macrophages and stimulates the release of inflammatory cytokines that contribute to

the further destruction of β-cells. IL-2 stimulates the differentiation and activation of $CD8^+$ precytotoxic T cells into cytotoxic effector T cells that destroy β-cells through the secretion of perforin and granzymes, as well as by Fas-mediated apoptosis of β-cells (Kukreja and Maclaren, 1999). In addition, the $CD4^+$ T cells can activate B cells to differentiate into plasma cells that produce antibodies specific to β-cell antigens such as insulin and GAD65. The antibodies bind to and initiate the destruction of β-cells via the complement system (Wallberg and Cooke, 2013). Immunoregulatory cytokines, IL-4 from Natural killer T (NKT) cells, transformation growth factor- β (TGF-β), and IL-10 from the forkhead box (Fox) p3+ regulatory T cells (Treg) inhibit the differentiation and function of $CD4^+$ T cells thereby conferring protection against the development of T1D (Wallberg and Cooke, 2013).

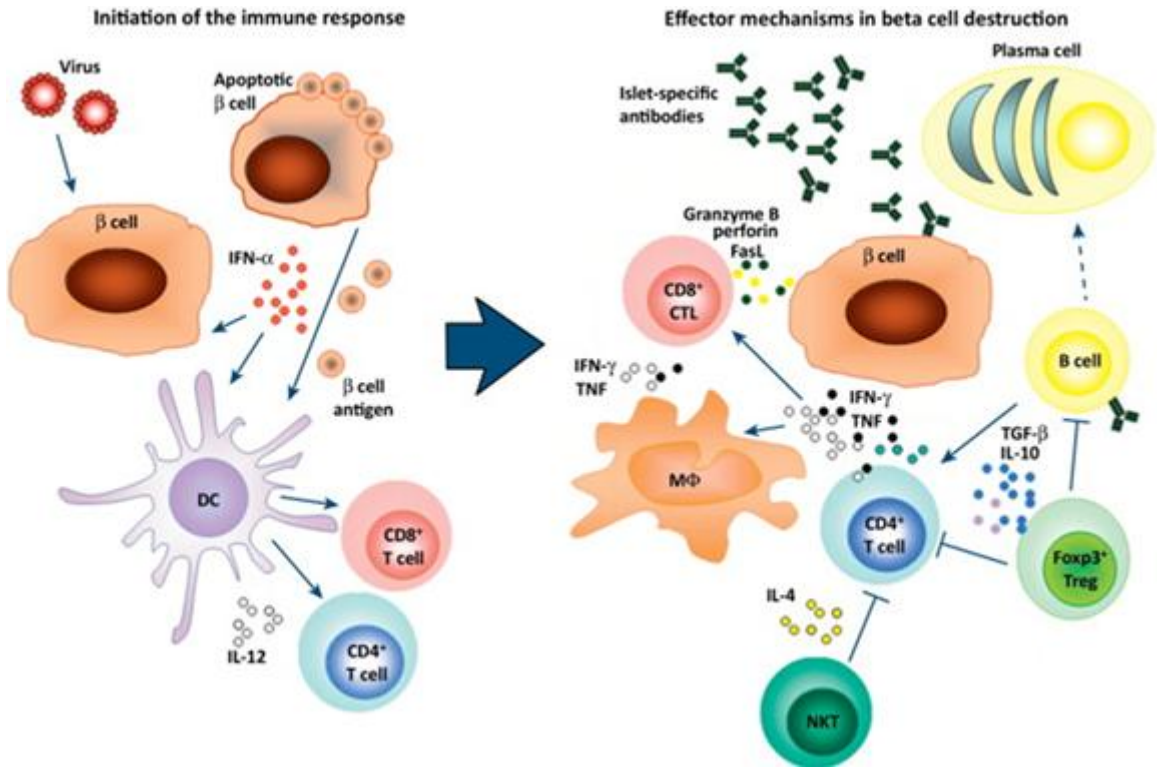


Figure 1.1: A schematic representation of the proposed pathogenesis of T1D (Wallberg and Cooke, 2013). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended).

1.2.2 Epidemiology of T1D

Type 1 diabetes can develop at any age, but it is most prevalent in children and young adults (Gerasimou *et al*. 2016). The peak age at diagnosis in European populations occurs between 10-14 years of age. However, a shift to a younger age at diagnosis has been reported with the largest increase in incidence seen in European individuals under the age of 5 years (Atkinson *et al.,* 2014). In contrast, in black South Africans, two peaks for the age at diagnosis of T1D have been observed; one at 11-15 years and a second peak at 26-30 years of age (Padoa *et al*., 2020).

In addition, the incidence of T1D varies geographically. Finland has the highest recorded incidence in the world of 62.3 per 100 000 inhabitants per year whereas China and Venezuela have the lowest incidence rate of 0.1 per 100 000 inhabitants per year (Verkauskiene *et al*., 2016; IDF, 2019). Studies in African countries investigating the incidence of T1D are scarce, however, they show that the incidence in sub-Saharan countries range from 1.5-10.1 cases per 100 000 (Padoa *et al*., 2020). It is likely that the numbers reported in Africa are an under-estimation (IDF, 2021).

1.3 Genetic susceptibility to T1D

Genome wide association studies (GWAS) have identified numerous T1D susceptibility genes such as the *human leukocyte antigen* (*HLA*) complex, *insulin (INS)*, *cytotoxic T lymphocyte antigen 4* (*CTLA4), and protein tyrosine phosphatase non-receptor type 22* (*PTPN22*) (Redondo *et al*., 2018). The majority of genes found to be associated with T1D play a role in immunity, insulin production and metabolism and protection against β-cell apoptosis (Atkinson, 2012).

1.3.1 The human leukocyte antigen complex and its role in T1D

The *HLA* complex is found on the short arm of chromosome 6 (Figure 1.2) and houses multiple genes that mediate adaptive immunity (Baschal *et al*., 2011). The *HLA* region consists of three classes of genes: class I, class II and class III. The *HLA* class II DRB1 and DQB1 genes were the first diabetes susceptibility genes to be identified and confer approximately 50% of the genetic risk for T1D (Balcha *et al.*, 2020). More recently, studies have shown that in addition to the *HLA* class II alleles, the class I alleles are similarly associated with a predisposition to T1D. Among these is *HLA-G* located at position 6p21.3 (Abdul-Hussein *et al.,* 2020).

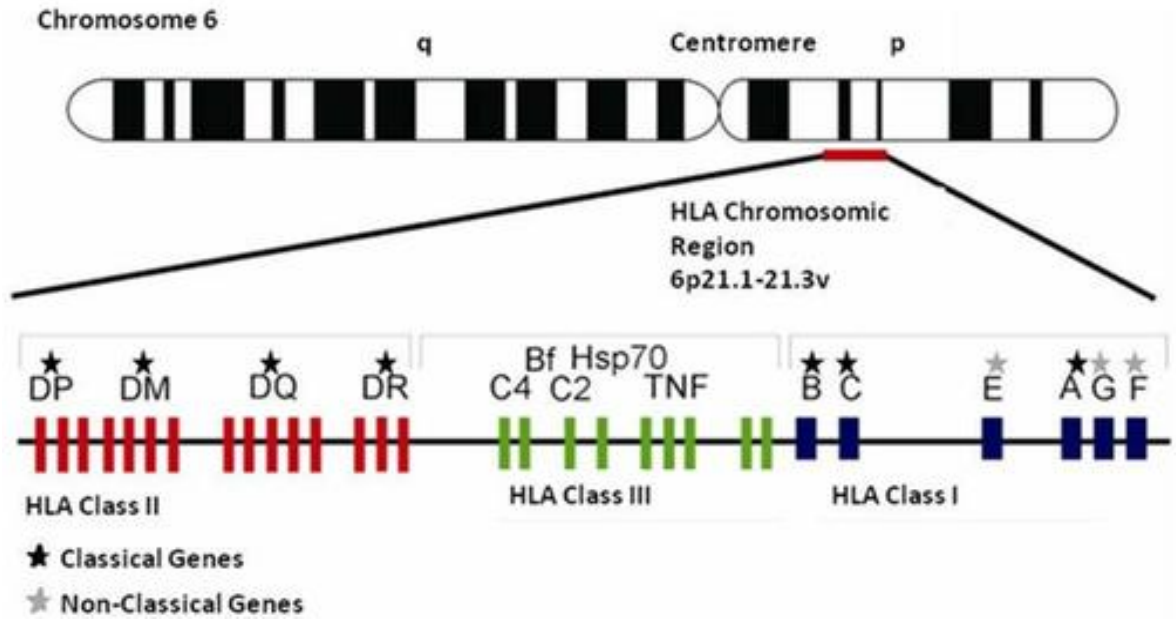


Figure 1.2: A schematic diagram of the position of the *HLA* region on the short arm of chromosome 6 including the classical and non-classical genes (Arnaiz-Villena *et al*., 2021). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended).

1.3.2 The human leukocyte antigen-G

HLA-G is classified as a non-classical *HLA* class Ib gene due to its limited allelic and protein variability, unique alternative splicing pattern that produces multiple isoforms, and immunosuppressive properties (Dahl *et al*., 2015; Xu *et al*., 2020). *HLA-G* is expressed in the placenta, cornea, thymus, pancreas, erythroid and endothelial cell precursors, and nail matrix (Rodrigues *et al*., 2020).

1.3.3 The *HLA-G* gene and protein structure

The *HLA-G* gene has seven introns and eight exons, each coding for a part of the heavy chain (Figure 1.3 A). Exon 1 codes for the signal peptide and exons 2-4 code for the extracellular domains α1, α2 and α3, respectively. Exons 5 and 6 code for the transmembrane and cytoplasmic domains, respectively (Donadi *et al*., 2011; Arnaiz-Villena *et al*., 2021). The HLA-G protein has a short cytoplasmic domain because of the presence of a premature stop codon in exon 6. Due to this stop codon, exon 7 is never present in the mature mRNA, and exon 8 forms part of the 3' untranslated region (UTR) which is involved in the transcriptional regulation of the gene (Carosella *et al*.,

2008). HLA-G exists in seven isoforms, four membrane-bound (HLA-G1 – HLA-G4) and three soluble (HLA-G5 – HLA-G7), due to alternative splicing of the primary *HLA-G* transcript (Figure 1.3 B; Rodrigues *et al*., 2020). HLA-G1 can also occur as a soluble molecule (sHLA-G1) as a result of proteolytic cleavage from the cell surface (Scarabel *et al*., 2020).

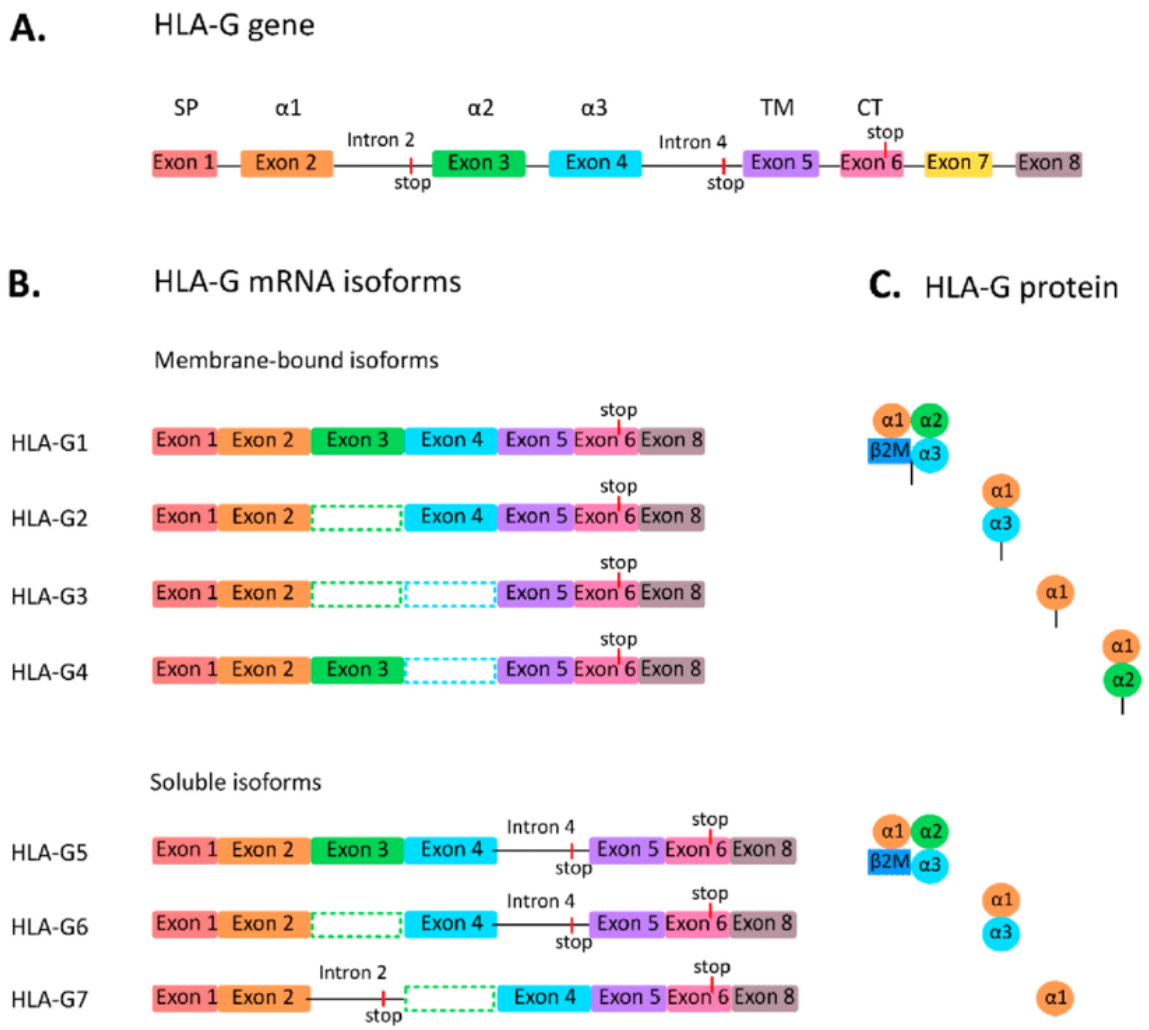


Figure 1.3: Schematic representation of the protein products coded by the *HLA-G* gene (Krijgsman *et al*., 2020). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended).

A: The *HLA*-G gene showing the position of the two stop codons relative to the exons. **B:** The seven HLA-G isoforms: HLA-G1 to HLA-G4 retain the transmembrane domain allowing them to be membrane bound. The transmembrane domain is absent in HLA-G5 to HLA-G6 due to the stop codon in intron 4 and in HLA-G7 due to the stop codon in intron 2 (van der Water *et al*., 2021). **C:** HLA-G protein isoforms: The full-length heavy chain is only present in HLA-G1

and HLA-G5 (van der Water *et al*., 2021). These two molecules form a heterodimer with a β2 microglobulin light chain by non-covalent binding to the α1 and α 3 domain (Attia *et al*., 2020).

1.3.4 The functions of HLA-G

The HLA-G molecule plays an important role in the suppression of immune responses, suggesting a role in T1D pathogenesis and/or disease progression (Arnaiz-Villena *et al.*, 2021). These immune-suppressive properties include altering cytokine secretion, downregulating the production of $CD4^+$ T cells, inhibiting the activity and inducing apoptosis of cytotoxic $CD8^+$ T cells and natural killer (NK) cells, inhibiting APCs and induction of regulatory T cells (Table 1.2) (Rizzo *et al*., 2014; Dahl *et al*., 2015; Wyatt *et al*., 2019). HLA-G isoforms exert their function mainly through direct interaction with specific inhibitory receptors on effector cells and indirectly through regulatory cells (Scarabel *et al*., 2020).

1.3.5 The role of HLA-G in T1D

The HLA-G molecule is constitutively expressed in the endocrine compartment of the pancreas, specifically in insulin secretory granules which contain multiple autoantigens (Cirulli *et al*., 2006). HLA-G is upregulated on the cell surface of islet cells that have been stimulated to secrete insulin (Cirulli *et al*., 2006). The activation of cytotoxic T cells relies on the density of autoantigens to HLA molecules on the cell surface thus, it has been suggested that local clustering of HLA-G during exocytosis of insulin may prevent deleterious T cell activation (Rezaei *et al*., 2019). When there are more HLA-G molecules on the cell-surface, T cell activation will be suppressed. Hence, a decrease in expression or production of HLA-G in the pancreas should be detrimental in individuals that are genetically prone to produce less HLA-G (de Albuquerque *et al*., 2016). It would lead to defects in tolerance to pancreatic autoantigens and therefore, contribute towards initiation and progression of T1D (Gerasimou *et al*., 2016).

Table 1.2: The different functions exerted by HLA-G upon interacting with cognate receptors that are expressed on a variety of immune cells

Target cell	HLA-G receptors	Function	References
Cytotoxic $CD8^+$ T cells	ILT2, KIR2DL4, CD8, CD160	Inhibits cytotoxic activity and induces apoptosis; inhibits chemotaxis	Carosella *et al.,* 2008; Chen *et al*., 2008; Xu *et al*., 2020
$CD4^+$ T cells	ILT2, CD160	Inhibits proliferation	Carosella *et al*., 2008; Chen *et al*., 2008
B cells	ILT2	Inhibits proliferation, differentiation, cytokine production, chemotaxis, and antibody secretion	Carosella *et al*., 2008; Martin-Villa *et al*., 2022
Natural killer cells	ILT2, KIR2DL4, CD160	Inhibits proliferation, cytotoxicity, and chemotaxis; induces apoptosis; upregulates inhibitory receptors; increases IFN y production	Carosella *et al*., 2008; Chen *et al*., 2008; Xu *et al*., 2020
Dendritic cells	ILT2, ILT4	Inhibition of DC maturation and antigen presentation	Carosella *et al*., 2008; Xu *et al*., 2020
Macrophages	ILT2, ILT4	Polarizes macrophages towards alternatively activated (M2) macrophages	Xu *et al*., 2020
Neutrophils	ILT4	Inhibits phagocytosis and the production of reactive oxygen species	Xu *et al*., 2020
DC-10	ILT2, ILT4	Increases expression of Tregs	Chen *et al*., 2008

DC-10: interleukin- 10 secreting dendritic cells; ILT2: Immunoglobulin-like transcription receptor type 2; ILT4: Immunoglobulin-like transcription receptor type 4; KIR2DL4: Killer immunoglobulin-like receptor 2DL4; Tregs: regulatory T cells.

1.4 *HLA-G* polymorphisms and T1D

A study of 2321 Norwegian families (with at least one family member having T1D) mapped the *HLA* region and identified the *HLA-G* gene as an independent locus for disease susceptibility (Eike *et al*., 2009). Polymorphisms in the *HLA-G* gene promoter modify the affinity of gene targeted sequences for transcription factors, altering the transcription levels of *HLA-G* (Alegre *et al*., 2014). Polymorphisms present in the *HLA-G* 3' UTR influence post-transcriptional regulation and consequently the mRNA stability. These polymorphisms may alter the target binding site of specific microRNAs (miRNA) to this region, modify polyadenylation signalling in the adenine and uridine (AU)-rich regulatory mRNA element and generate alternative splice sites (Castelli *et al*., 2014). In addition, the presence of polymorphisms within exons may lead to changes in the molecule's amino acid sequence which can alter production of HLA-G isoforms, peptide binding and immunoregulatory functions (Donadi *et al*., 2011). Therefore, some of these polymorphisms may subsequently be involved in the initiation and/or progression of disease. However, there is little data regarding the association of *HLA-G* gene polymorphisms with T1D.

1.4.1 The HLA-G 14 bp insertion/deletion polymorphism

A 14 bp (5′-ATTTGTTCATGCCT-3') insertion/deletion (indel) polymorphism in exon 8 of the *HLA-G* gene (Harrison *et al*., 1993), has been associated with both an increased risk of developing T1D and a younger age at diagnosis of the disease (Silva *et al*., 2016). A study by Rezaei *et al*. (2019) found that the *HLA-G* 14 bp homozygous insertion genotype conferred a 3.82-fold increased risk of developing T1D in the Iranian population. In contrast, Silva *et al.* (2016) found that in the Brazilian population, the deletion allele and the homozygous deletion genotype were associated with susceptibility to T1D. A study conducted in a Cypriot population, found no significant difference in the frequency of the 14 bp indel polymorphism between participants with T1D and controls. However, the homozygous deletion genotype was associated with an earlier age of onset amongst the participants with T1D (Geramisou *et al*., 2016). The frequency of the homozygous deletion genotype was almost threefold higher in early onset participants with T1D (0-8 years of age) compared to those with a late onset (13-39 years of age). This suggests that the deletion genotype is associated with an earlier presentation of autoimmune diabetes and indicates a failure of HLA-G to confer sufficient immune-tolerance at the pancreatic level (Geramisou *et al*., 2016).

The presence of the 14 bp insertion generates an additional splice site resulting in the removal of 92 bases from the start of exon 8 which results in a truncated mRNA that is more stable than the mRNA transcript generated by the 14 bp deletion (Rousseau *et al*., 2003; Castelli *et al*., 2014). However, the insertion allele has been associated with lower serum and plasma levels of membrane bound and soluble HLA-G (Rousseau *et al*., 2003). This may be due to the finding that the 14 bp insertion mRNA transcript is expressed at significantly lower levels (Figure 1.4) than the *HLA-G* mRNA isoform in which the 14 bp sequence is deleted (Rousseau *et al*., 2003). The higher stability conferred by alternative splicing does not compensate for the lower *HLA-G* expression associated with the 14 bp insertion sequence (Castelli *et al*., 2014).

The *HLA-G* insertion is found in linkage disequilibrium with the rs1063320 +3142 G allele (Dahl *et al*., 2015). The presence of a G at position +3142 increases the affinity of the hsa-miR-148a, hsa-miR-148b, and hsa-miR-152 miRNAs for the 3' UTR of the *HLA-G* mRNA, leading to mRNA degradation and downregulation of *HLA-G* expression (Castelli *et al*., 2009). Therefore, there is a decrease in mRNA availability and hence the production of HLA-G.

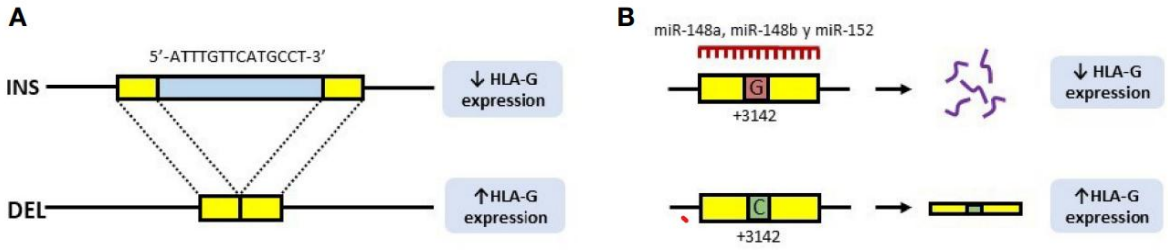


Figure 1.4: Schematic representation of the 14 bp *HLA-G* indel and its effect on *HLA-G* expression (Martin-Villa *et al.,* 2022*).* This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended).

A. The presence of the 14 bp insertion allele influences mRNA stability and has been associated with decreased *HLA-G* expression. **B:** The +3142 G allele increases the affinity for miRNAs which subsequently increases mRNA degradation and reduces HLA-G availability.

1.4.2 The *HLA*-G rs9380142 (+3187 A>G) polymorphism

The rs9380142 single nucleotide polymorphism (SNP) occurs at position +3187 in the 3' UTR of the *HLA-G* gene. It results in an adenine to guanine change and influences both *HLA-G* mRNA stability and availability for translation (Castelli *et al*., 2009). The SNP is near an AU-rich motif (Figure 1.5) which acts as a binding site for proteins that mediate mRNA degradation (Catelli *et al*., 2014). The presence of the A allele increases the number of adenines in this motif which decreases *HLA-G* mRNA stability (Alegre *et al*., 2014). Reduced mRNA stability results in its degradation, lower mRNA levels and consequently, lower serum/plasma levels of HLA-G. This polymorphism's association with T1D is however, unknown.

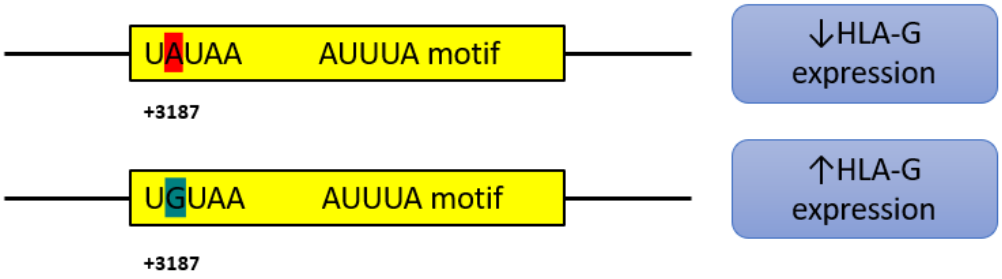


Figure 1.5: The effect of the HLA-G +3187 (A>G) SNP on the expression of the *HLA-G* gene (adapted from Martin-Villa *et al.,* 2022). This figure is copied in terms of Section 12 of the Copyright Act No 98 of 1978 (amended).

The occurrence of the +3187 A allele has been associated with decreased mRNA stability and low HLA-G production due to its vicinity to an AU-rich motif.

1.5 Aim and study objectives

Literature suggests an important role of *HLA-G* in the pathogenesis of T1D and age at diagnosis. However, to our knowledge, there are no studies investigating these associations in the South African black population. Therefore, the aim of this study was to determine whether polymorphisms in the *HLA-G* gene are associated with age at diagnosis and/or T1D in the South African black population. In addition, this study aimed to determine the association of *HLA-G* genotypes with serum HLA-G concentrations. The objectives of the study were as follows:

1) To sequence the *HLA-G* gene in 20 participants with T1D (cases) and 20 participants without diabetes (controls) to identify novel variants that may be associated with T1D
2) To genotype cases and controls for a maximum of two variants identified from sequencing analysis as having the greatest frequency differences between cases and controls
3) To genotype cases and controls for the *HLA-G* 14 bp indel and the +3187A>G (rs9380142) polymorphisms
4) To measure serum concentrations of HLA-G in cases and controls
5) To compare genotypic frequencies of the *HLA-G* gene polymorphisms between cases and controls and relate these to serum HLA-G concentrations and age at diagnosis

2. Methods

Recipes for buffers used in the study can be found in Appendix A.

2.1 Study participants

Black South Africans with T1D (cases; n=263) were recruited from the diabetes clinics at Charlotte Maxeke Johannesburg Academic Hospital, Chris Hani Baragwanath Academic Hospital and Steve Biko Academic Hospital from November 2007 to December 2012 as well as from October 2014 to September 2015. Participants were classified as having T1D based on clinical criteria: age at diagnosis ≤30 years and insulin therapy within the first year of diagnosis (therapy initiated immediately in children). Participants older than 30 years of age at diagnosis but requiring insulin treatment within the first year of diagnosis and who were not obese were classified as having T1D. Healthy control participants (n=202) were recruited between October 2014 to September 2015 from Wits Medical School and South African National Blood Services (SANBS) blood drives. Participants with clinical evidence of chronic pancreatitis, type 2 diabetes or gestational diabetes were excluded from the study. In addition, control participants with blood glucose levels >11.1 mmol/L and < 18 years of age were excluded from the study. Biographic (age, gender), anthropometric (BMI, waist and hip circumference), and clinical (HbA1c, glucose and blood pressure) data were available for all participants. HbA1c was not measured in the control group.

Ethics clearance was obtained from the Human Research Ethics Committee (Medical) of the University of the Witwatersrand (ethics clearance number M200174 and M220484) and the SANBS Research Ethics Committee (clearance certificate number 2014/19) (Appendix B).

2.1.1 Sample size

The sample size was based on two previous studies. A Brazilian study showed an association between the *HLA-G* 14 bp indel polymorphism and T1D in 81 cases and 100 controls (Silva *et al.,* 2016). Similarly, in a study with a sample size of 170 participants with T1D, an association was observed between the 14 bp indel polymorphism and age at diagnosis in a Greek population (Geramisou *et al*., 2016). The frequencies of alleles for the *HLA-G* 14 bp indel polymorphism in the 1000 genomes project database were similar in the American (0.62), European (0.63) and

African (0.62) populations (The 1000 Genomes Project Consortium, 2015). Therefore, we expect to be able to see associations between genotype and T1D and/or age at diagnosis with a similar number of participants in the South African black population.

2.2 Sequencing of the *HLA-G* gene

To identify *HLA-G* gene variants that are potentially associated with T1D, all eight exons of the gene were sequenced, and the nucleotide sequences were compared in 20 cases and 20 controls.

2.2.1 PCR amplification

Primers flanking the exons were designed using the primer3 software (Untergasser *et al.,* 2012) and synthesised by Inqaba Biotech, Pretoria, South Africa (Table 2.1).

Table 2.1: Primer sequences flanking exons of the *HLA-G* gene and the corresponding PCR amplicon fragment sizes

Primers	Primer sequence 5' – 3'	Fragment sizes (bp)
Exon 1 - 2		
Forward primer	GCGGTCCTGGTTCTAAAGTC	566
Reverse primer	TCGTGATCTGCTCCCTGG	
Exon 3		
Forward primer	AGGCGCCTTTACCAAAATCC	399
Reverse primer	CTACAGGAGATCAGGGAGGC	
Exon 4 - 5		
Forward primer	AGTGTCCCATGAGAGATGCA	638
Reverse primer	ACTTCTACCTGGGGCTTGAA	
Exon 6 - 8		
Forward primer	ACTTCTCGAGGGTCCAAGAC	840
Reverse primer	GTGGAAAGTTCTCATGTCTTCCA	

PCR reactions were set up in 0.2 ml microcentrifuge tubes in a final volume of 20 µl consisting of 10 µl of Q5® High-Fidelity 2X Master Mix (New England Biolabs® Inc., Massachusetts, USA), 0.25 µl of the forward and reverse primer (10µM), 7.5 µl of deionised water (diH_2O), and 2 µl of DNA (~40ng). A non-template control (NTC) was included in each PCR run to check for contamination. PCR reactions were carried out in a T100 thermal cycler (Bio-Rad Laboratories, California, USA). The PCR conditions were as follows: initial denaturation at 98°C for 10 seconds, 30 cycles of denaturation (98°C for 30 seconds), annealing (65°C for 30 seconds), and extension (72°C for 20 seconds) with a final extension at 72°C for two minutes.

2.2.2 Agarose gel electrophoresis

To confirm if the correctly sized amplicon was generated, agarose gel electrophoresis was performed. A 1 % agarose gel was made by dissolving 1 g of SeaKem LE agarose powder (Lonza, Basel, Switzerland) in 100 ml of 1X Tris–Borate-EDTA (TBE) buffer and microwaving the solution for two minutes. Following the addition of ethidium bromide (1 µl; 10 mg/ml; Sigma-Aldrich, Missouri, USA) to the mixture, the gel was poured into a casting tray and allowed to set at room temperature. PCR products (3 µl) were mixed with 2 µl of loading dye and loaded into the wells of the gel. A 100 bp DNA ladder (1 µl; Solis BioDyne, Tartu, Estonia) was included in one of the wells for size comparison. The gel was run for 30 minutes at 100 volts (V) and visualized under UV light in a ChemiDoc UV transilluminator (Bio-Rad, California, USA).

The PCR products for exon 4-5 and exon 6-8 had spurious bands despite multiple attempts at optimising the PCR reaction and therefore these products needed to be gel purified.

2.2.3 DNA purification for exon 4-5 and exon 6-8

PCR amplicons of the correct size for exon 4-5 and exon 6-8 were excised from a 1% agarose gel and purified using the ZymoClean™ Gel DNA Recovery kit (ZYMO RESEARCH, California, USA) according to the manufacturer's instructions. The centrifugation steps described below were done using the Mikro 200 Hettich microcentrifuge (Hettich, Tuttlingen, Germany).

Excised amplicons were placed in 1.5 ml microcentrifuge tubes and the weight of the gel slice determined. Three volumes of Agarose Dissolving Buffer (ADB) for each mg of gel slice were added to the microcentrifuge tube (e.g., for 100 mg excised gel add 300 µl buffer). The tubes were incubated at 55˚C for 10 minutes in a heating block (Lab-Line Multi Blok Heater, Mumbai, India) until the agarose was completely melted. The melted agarose solutions were transferred into Zymo-Spin™ columns assembled in collection tubes and centrifuged for one minute at 11 000 *x g*. The flowthrough was discarded, DNA wash buffer (200 µl) added, and the columns centrifuged for 30 seconds at 11 000 *x g*. The wash step was repeated and thereafter, the columns were placed in a clean 1.5 ml microcentrifuge tube. Elution buffer (10 µl) was added to each column and the tubes centrifuged for one minute at 11 000 *x g* to elute the purified product.

The eluted PCR product (2 µl) was run on a 1% agarose gel for 30 minutes at 100 V to ensure there was adequate recovery of the amplicon. The DNA fragments were visualized under UV light in a ChemiDoc UV transilluminator.

2.2.4 Sequencing of PCR amplicons

The PCR products for exons 1 to 8 from the 40 participants (section 2.2) were sent to Inqaba Biotechnology for Sanger sequencing in the forward and reverse direction using the primers listed in Table 2.1.

2.2.4.1 Identification of *HLA-G* variants

The sequencing data was viewed using Sequencher version 5.4.6 software (Gene Codes Corporation, Michigan, USA). *HLA-G* exon sequences were aligned with the reference sequence (GRCh38.p14; NC_000006.12) to identify variants (Appendix C). All variants were compared to reported polymorphisms in the NCBI SNP database variation viewer (http://www.ncbi.nlm.nih.gov/variation/view). Two variants in exon 8, rs1063320 and rs1710, which had the highest allele frequency differences between cases and controls were selected for further analysis.

2.3 Genotyping participants for the *HLA-G* rs1063320 and rs1710 polymorphisms

2.3.1 PCR amplification

Participants were genotyped for the rs1063320 and rs1710 polymorphisms by PCR-restriction fragment length polymorphism (RFLP). A 255 bp fragment of the *HLA-G* gene flanking both polymorphisms was amplified using primers designed with primer3 software (Table 2.2) (Untergasser *et al.,* 2012).

Table 2.2: Primer sequences flanking the region containing the HLA-G rs1063320 and rs1710 polymorphisms

Primer	Primer sequence 5' – 3'	Manufacturer
Forward primer	CCCTTTGTGACTTCAAGAACCC	Inqaba Biotech, Pretoria,
Reverse primer	ACCCATCAATCTCTCTTGGAAA	South Africa

PCR reactions were set up in 0.2 ml microcentrifuge tubes in a final volume of 25 µl and consisted of 12.5 µl of One*Taq*® 2X Master Mix (New England Biolabs® Inc., Massachusetts, USA), 1 µl of both the forward and reverse primer (10µM), 8.5 µl of diH_2O, and 2 µl of DNA. The final primer concentration was 0.4 µM. An NTC was included in each PCR run to check for contamination. PCR reactions were carried out in a T100 thermal cycler. The PCR conditions were as follows: initial denaturation at 95 °C for three minutes, 30 cycles of denaturation (95 °C for 30 seconds), annealing (56.1°C for 30 seconds), and extension (72 °C for 30 seconds) with a final extension at 72 °C for five minutes.

2.3.2 Agarose gel electrophoresis

PCR products (4 µl) were mixed with 1 µl loading dye and run on a 1 % agarose gel together with 1 µl of a 100 bp DNA ladder (Solis BioDyne) at 100 V for 30 minutes. The amplified PCR products were visualized under UV light in a ChemiDoc UV transilluminator.

2.3.3 Restriction digest

PCR products were digested with ApaLI and PstI (New England Biolabs Inc., Massachusetts, USA) to genotype participants for the rs1063320 and rs1710 polymorphisms, respectively. ApaLI recognises the GTGCAC cut site and cuts when the rs1063320 G allele is present. PstI recognises the CTGCAG cut site and cuts when the C allele of rs1710 is present. PCR products were digested in a final volume of 20 µl using the reagents listed in Table 2.3. Samples were incubated for 16 hours at 37°C followed by a 20-minute incubation at 80°C to inactivate the enzyme for 30 minutes at 37°C for ApaLI and PstI and two hours.

Table 2.3: Reagent volumes for the digestion of rs1063320 and rs1710 PCR products

	***HLA-G* single nucleotide polymorphisms**	
Reagent	**rs1063320**	**rs1710**
PCR product	10 µl	10 µl
rCutsmart Buffer™ (10X)	2µl	-
NEBuffer™ r3.1 (10X)	-	2 µl
ApaLI (10 000 U/ml)	0.5 µl	-
PstI (20 000 U/ml)	-	0.5 µl
diH_2O	7.5 µl	7.5 µl

The digested products for ApaLI and PstI (7 µl) were mixed with 2 µl of loading dye and run on a 2 % agarose gel at 100 V for 60 minutes together with 1 µl of a 100 bp DNA ladder (Solis BioDyne) and an undigested PCR product (negative control). Gels were visualized under UV light in a ChemiDoc UV transilluminator and participant genotypes were determined based on the fragment sizes generated (Table 2.4).

Table 2.4: PCR-RFLP fragment sizes and their corresponding genotypes for the HLA-G rs1063320 and rs1710 polymorphisms

Genotype	Fragment sizes (bp)
rs1063320	
CC	255
CG	255, 165 and 90
GG	165 and 90
rs1710	
GG	255
GC	255, 216 and 39
CC	216 and 39

2.4 Genotyping participants for the *HLA-G* 14 bp I/D polymorphism

2.4.1 PCR amplification

Participants were genotyped for the presence of the *HLA-G* 14 bp indel by conventional PCR. A 180 bp fragment in the 3' UTR of the *HLA-G* gene was amplified using primers (Table 2.5) designed by primer3 software (Untergasser *et al.,* 2012) to flank the indel.

Table 2.5: Primer sequences flanking the region containing the 14 bp indel

Primer	Primer sequence 5' – 3'	Manufacturer
Forward primer	CTGTGTGGGACTGAGTGGC	Inqaba Biotech, Pretoria, South Africa
Reverse primer	GAGACATCCCAGCCCCTTTT	

PCR reactions were set up in 0.2 ml microcentrifuge tubes in a final volume of 20 µl and consisted of 10 µl of One*Taq*® 2X Master Mix, 1 µl of each of the forward and reverse primers (10µM), 6 µl of diH_2O, and 2 µl of DNA. An NTC was included in each PCR run to check for contamination. PCR reactions were carried out in a T100 thermal cycler, and the conditions were as follows: initial denaturation at 95°C for three minutes, 30 cycles of denaturation (95°C for 30 seconds), annealing (65°C for 30 seconds), and extension (72°C for 30 seconds) with a final extension at 72°C for five minutes.

2.4.2 Agarose gel electrophoresis

PCR products (6 µl) were mixed with 2 µl of loading dye and loaded into the wells of a 3 % agarose gel. A 100 bp DNA ladder (1 µl; Solis BioDyne) was included in one of the wells. The gel was run at 100 V for 90 minutes and PCR amplicons were visualized under UV light in a ChemiDoc UV transilluminator. The genotypes were determined based on the resultant PCR product sizes (Table 2.6).

Table 2.6: PCR amplicon fragment sizes and their corresponding genotypes for the 14 bp indel

Genotype	Fragment sizes (bp)
Deletion/Deletion	166
Deletion/Insertion	166 and 180
Insertion/Insertion	180

2.5 Genotyping participants for the *HLA-G* rs9380142 polymorphism

Participants were genotyped for the rs9380142 polymorphism by real-time PCR using a predesigned Applied Biosystems TaqMan® SNP Genotyping assay (Massachusetts, USA)**.** The TaqMan® Assay contained the forward and reverse primers, and two TaqMan® probes labelled with VIC™ and FAM™ dyes for detection of the A and G alleles, respectively. TaqMan® probes are sequence specific oligonucleotides with a fluorophore at the 5' end and a quencher at the 3' end. When the probes are intact, the quencher, due to its proximity to the fluorophore, is able to absorb the light emitted by the fluorophore. During the PCR, the probes bind to their complementary sequence on the template. As Taq polymerase extends from the primer, the enzyme will degrade the probe, upon contact using its 5'-3' exonuclease activity and releases the fluorophore. This allows for the fluorescent signal to be detected and for genotypes to be determined (Jothikumar *et al*., 2009).

2.5.1 Real-time PCR amplification

Prior to setting up the PCR run, the 40X TaqMan® SNP Genotyping assay was diluted to a 20X working solution with 1X Tris-EDTA (TE) buffer. The PCR reactions were prepared in a final volume of 25 µl consisting of 12.5 µl of the 2X TaqMan® mastermix, 1.25 µl of the 20X SNP Genotyping assay, 9.25 µl of RNase free water and 2 µl of

DNA. An NTC was included in each PCR run to check for contamination. PCR reactions were carried out in a Rotor-Gene™ 6000 Real-time PCR thermocycler (Corbett Research, Sydney, Australia).

The PCR conditions were as follows: polymerase activation at 95°C for 10 minutes and 35 cycles of denaturation at 95°C for 15 seconds and a combined annealing and extension step at 60°C for 1 minute. The A allele of rs9380142 was detected with a VIC® dye labelled probe and the G allele was detected with a FAM™ labelled probe. The Rotor-Gene Q software version 2.3.5 was used to genotype participants with the rs9380142 polymorphism using an allelic discrimination plot.

2.6 Confirmation of genotypes by sequencing

PCR products representing each genotype for the rs1063320, rs1710 and the 14 bp indel were sent to Inqaba Biotech (Pretoria, South Africa) for sequencing to confirm the genotypes obtained. The sequencing data was viewed using Sequencher version 5.4.6 software. The rs9380142 SNP was not sequenced as it was an optimised custom assay.

2.7 Measurement of serum HLA-G concentrations

The concentration of HLA-G was measured in serum using the Elabscience® Human MHCG/HLA-G (Major Histocompatibility Complex Class I G) ELISA kit according to the manufacturer's instructions (Elabscience Biotechnology Inc., Houston, USA). The reagents and samples were equilibrated to room temperature prior to use. All centrifugation steps described below were performed using the Mikro 200 Hettich microcentrifuge (Hettich, Tuttlingen, Germany).

2.7.1 Preparation of reagents and participant serum samples

A 1x wash buffer was prepared by adding 720 ml of diH_2O to 30 ml of the 25x Concentrated Wash Buffer. The 100x Biotinylated Detection Ag/Ab was centrifuged at 800 *x g* for one minute and 120 µl added to 11 880 µl of the Biotinylated Detection Ab Diluent to make a 1x working solution. The 100x Concentrated Horseradish peroxidase (HRP) Conjugate was centrifuged for one minute at 800 *x g* and diluted to a 1x working solution by adding 120 µl of the concentrate to 11 880 µl of the HRP Conjugate Diluent. The positive (12.45 - 20.75 ng/ml) and negative (0 ng/ml) controls

were reconstituted by adding 1 ml of Reference Standard and Sample (RSS) diluent. They were allowed to stand for 10 minutes with intermittent inversion. The participant serum samples were diluted 5x in Reference and Sample diluent.

2.7.2 Preparation of HLA-G standards

The HLA-G reference standard was centrifuged for one minute at 10 000 *x g* and reconstituted in 1 ml of RSS diluent to produce the 40 ng/ml standard. The solution was allowed to stand for 10 minutes with intermittent inversion. Eight standards were prepared, by serial dilution (Figure 2.1), with concentrations ranging from 0 to 40 ng/ml. All standards were assayed in duplicate.

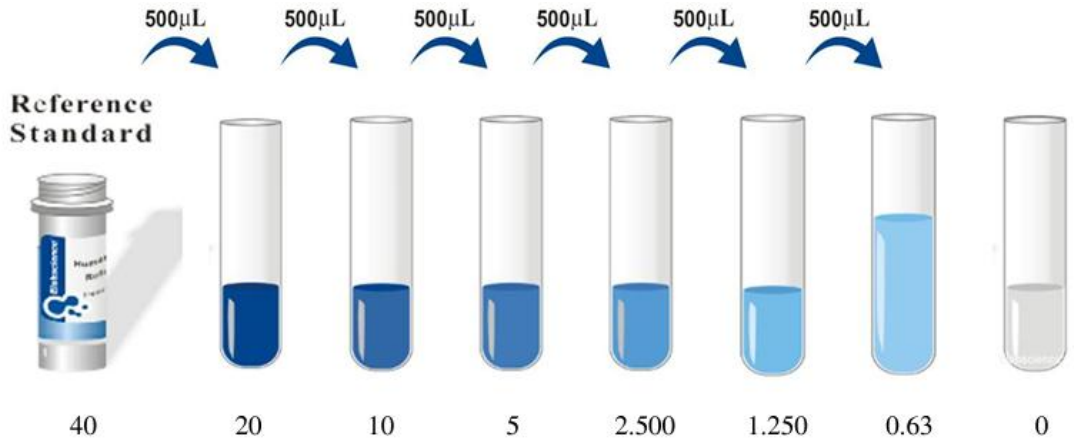


Figure 2.1: Dilution series of HLA-G standards

2.7.3 Assay procedure

Standards, controls, and serum samples (100 µl) were added to the assigned wells of a 96-well plate. The plate was sealed and incubated for 90 minutes at 37°C in a Thermostar orbital shaker (BMG Labtech, Ortenberg, Germany). The liquid was decanted from the plate and 100 µl of Biotinylated Detection Ab/Ag working solution (1x) was added to each well. The plate was sealed and incubated for one hour at 37°C in the orbital shaker. The liquid was decanted, and the plate washed three times with 350µl of 1x wash buffer using the Biotek ELX50 Microplate Washer (Agilent Technologies, California, USA).

Following the last wash, 100 µl of HRP Conjugate working solution (1x) was added to each well, the plate was sealed and incubated for 30 minutes at 37°C in the orbital shaker. The solution was decanted, and the plate washed five times as above. Thereafter, 90 µl of Substrate Reagent was added to each well, the plate sealed and incubated at 37°C for 15 minutes in the orbital shaker. Finally, 50 µl of Stop solution was added to each well and the optical density (OD) was determined at 450 nm using the Synergy HT plate reader (BioTek Instruments Inc., Vermont, USA) that had been preheated for 15 minutes. The concentrations of serum HLA-G were determined from the standard curve (Appendix D) using the Gen5 version 1.11.5.0 microplate reader and imager software (BioTeck Instruments Inc., Vermont, USA).

2.8 Statistical analysis

Statistica version 14.0 software (TIBCO Software Inc., California, USA) was used for all statistical analyses. Continuous variables were determined to be normally distributed if their skewness values fell between -1 and 1. The duration of disease, blood glucose concentrations and serum HLA-G concentrations were log transformed to normality. The parametric data were reported as mean ± standard deviation (SD) and non-parametric data represented as median [lower quartile; upper quartile]. Categorical variables were analysed using a Chi square test and values were presented as frequencies or percentages.

Due to the small number of participants with the minor genotypes, these were combined with heterozygous genotypes to increase numbers and allow for statistical analysis. A students t-test was used to compare normally distributed continuous variables between two groups and an analysis of variance (ANOVA) test was used when comparing more than two groups. A multiple linear regression model was used to determine factors affecting systolic blood pressure (SBP) when controlling for T1D status, age, BMI, sex, and *HLA-G* rs1063320, rs1710 and 14 bp indel polymorphisms. A logistic regression model was performed to ascertain the variables associated with the risk of developing T1D controlling for sex, BMI, *HLA-G* rs1063320, rs1710 and 14 bp indel polymorphisms and serum HLA-G concentration. To ascertain the determinants of age at diagnosis of T1D in our study population, a logistic regression model was performed, controlling for sex and the rs1063320, rs1710 and 14 bp indel polymorphisms. The rs9380142 genotypes were not included in the regression models

due to the small number of participants genotyped for this polymorphism. Furthermore, Michael H. Court's (2005–2008) Excel-based Hardy-Weinberg equilibrium (HWE) calculator was used to determine if the polymorphisms were in HWE.

3. Results

The aim of this study was to determine whether polymorphisms in the *HLA-G* gene are associated with age at diagnosis and/or T1D in the South African black population. In addition, this study also aimed to determine the effect of *HLA-G* genotypes on serum HLA-G concentrations.

3.1 Anthropometric and clinical characteristics of the study population

A total of 465 black South Africans were previously recruited for this study. The comparison of anthropometric and clinical characteristics between cases and controls is presented in Table 3.1. The cases and controls were matched for age (p=0.823) and sex (p=0.398). The mean age at diagnosis of cases was 19.3 ± 10.1 years and median duration of disease was 6.00 [2.00; 12.0] years.

The control participants had a significantly higher BMI (27.0 ± 6.1 vs. 25.1 ± 6.77 kg/m^2; p<0.001). Blood glucose concentrations (8.20 [5.10;13.5] vs. 5.40 [6.40;7.10] mmol/L; p<0.001), SBP (128.0 ±17.9 vs. 121.0 mmHg; p=0.021) and serum HLA-G concentrations (52.0 [32.0; 89.7] vs. 32.8 [22.0; 51.9] ng/ml; p<0.001) were significantly higher in cases than controls.

Table 3.1: Comparison of anthropometric and clinical characteristics between the cases and controls groups

Variable	Cases (n=263)	Controls (n=202)	p value
Age (years)	27.3 ± 12.0	27.1 ± 8.76	0.843
Age at diagnosis (years)	19.3 ± 10.1	-	-
Duration of disease (years)	6.00 [2.00; 12.0]	-	-
Sex			
Male (%, n)	46.0 (121)	42.1(85)	0.398
Female (%, n)	54.0 (142)	57.9 (117)	
BMI (kg/m^2)	25.1 ± 6.77	27.1 ± 6.11	0.001
HbA1c (%)	10.3 ± 2.80	-	-
Glucose (mmol/L)	8.20 [5.10; 13.5]	5.40 [4.60; 7.15]	<0.001
SBP (mmHg)	127.7 ± 17.9	121.0 ± 11.5	0.021
DBP (mmHg)	78.3 ± 11.9	77.6 ± 9.85	0.707
HLA-G conc (ng/ml)	52.0 [32.0; 89.7]	33.1 [22.0; 52.0]	<0.001

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, HbA1c = glycated haemoglobin, SBP = systolic blood pressure, DBP = diastolic blood pressure, HLA-G= human leukocyte antigen-G

3.2 Sequencing of *HLA-G* exons

To identify *HLA-G* variants that may be associated with T1D, all eight exons of the gene were sequenced in 20 cases and 20 controls. A total of 103 variants were identified (Appendix C). Of these, 70 were known variants and 33 were novel. Table 3.2 presents the top 10 variants with the highest genotypic frequency difference between cases and controls. Two 3' UTR variants, rs1063320 and rs1710 were selected for further analysis.

Table 3.2: *HLA-G* variants and their frequencies in cases and controls

Variant location	Allele	Frequencies (n)	
		Cases	Controls
Exon 1			
rs1630223 (G>A)	G	0.556 (10)	0.833 (15)
	A	0.444 (8)	0.167 (3)
rs1630185 (G>A)	G	0.556 (10)	0.833 (15)
	A	0.444 (8)	0.167 (3)
Intron 1-2			
rs1629329 (T>C)	T	0.528 (19)	0.824 (14)
	C	0.472 (17)	0.176 (3)
rs1628628 (C>T)	C	0.528 (19)	0.824 (14)
	T	0.472 (17)	0.176 (3)
Exon 2			
rs1130355 (G>A)	G	0.579 (11)	0.833 (15)
	A	0.421 (8)	0.167 (3)
Exon 4			
rs878913602 (G>A)	G	0.562 (9)	0.227 (5)
	A	0.438 (7)	0.773 (17)
Exon 5			
rs747314068 (G>A)	G	0.438 (9)	0.615 (8)
	A	0.562 (7)	0.385 (5)
Intron 7-8			
rs915668 (G>C)	G	0.667 (12)	0.423 (11)
	C	0.333 (6)	0.577 (5)
Exon 8			
rs1710 (G>C)	G	0.333 (1)	0.75 (9)
	C	0.667 (2)	0.25 (3)
rs1063320 (C>G)	C	0.333 (1)	0.75 (9)
	G	0.667 (2)	0.25 (3)

3.3 Determining the *HLA-G* gene genotypes for rs1063320, rs1710, 14 bp indel and rs9380142

3.3.1 Detection of the *HLA*-G rs1063320and rs1710 SNPs

The genotypes of the rs1063320 C>G and rs1710 G>C polymorphisms were determined by PCR-RFLP. A 255 bp region of the *HLA-G* gene 3' UTR containing both polymorphisms was amplified by PCR (Figure 3.1). The PCR amplicons for rs1063320 and rs1710 were digested with the ApaLI (Figure 3.2) and PstI (Figure 3.3) restriction enzymes, respectively.

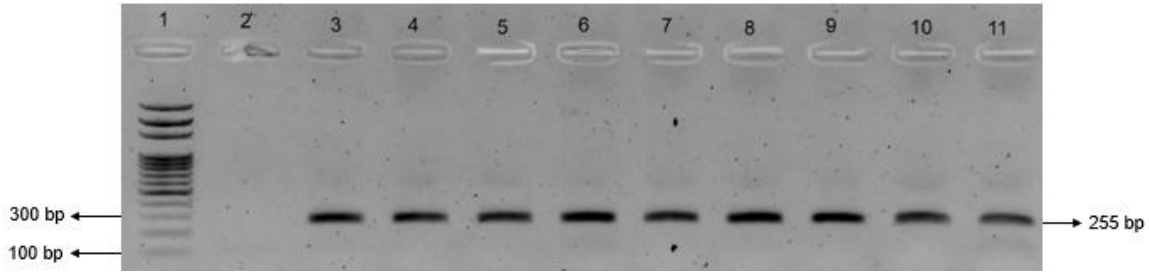


Figure 3.1: A 1 % agarose gel showing the PCR products obtained from the amplification of a 255 bp region of the *HLA-G* gene containing the rs1063320 and rs1710 polymorphisms

Lane 1: 100 bp DNA ladder; Lane 2: NTC; Lane 3-11: PCR amplicons with a fragment size of 255 bp.

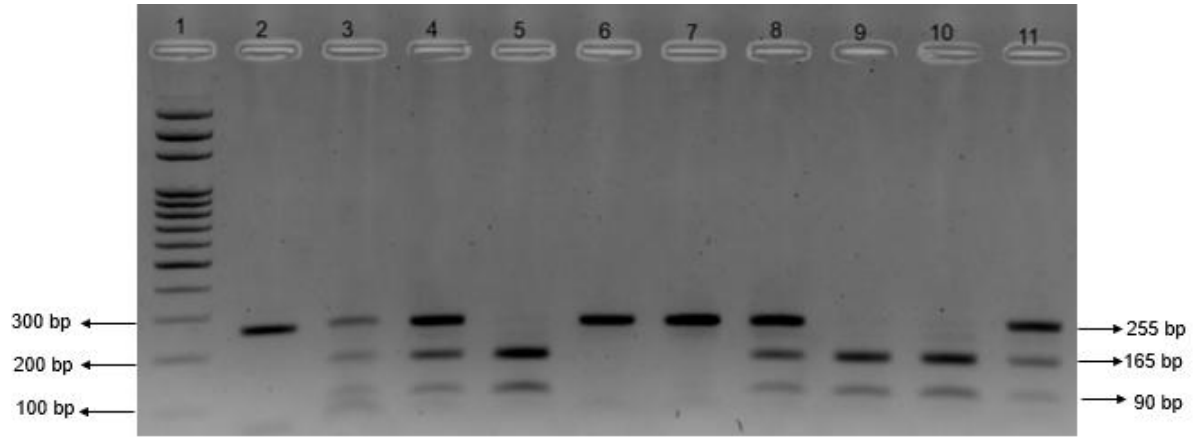


Figure 3.2: A 2% agarose gel showing the restriction fragment profile of the *HLA-G* rs1063320 polymorphism PCR products digested with ApaLI

Lane 1: 100 bp DNA ladder; Lane 2: undigested PCR amplicon (255 bp); Lane 3-4, 8 and 11: participants heterozygous for the rs1063320 SNP (CG genotype; fragment sizes: 255, 165 and 90); Lane 5, 9 and 10: participants homozygous for the rs1063320 G allele (GG genotype; fragment sizes: 165 and 90 bp); Lane 6 and 7: participants homozygous for the rs1063320 C allele (CC genotype; fragment size: 255 bp). All bands below the 100 bp ladder marker represent primer dimers which result from an unintended dimerization of primers.

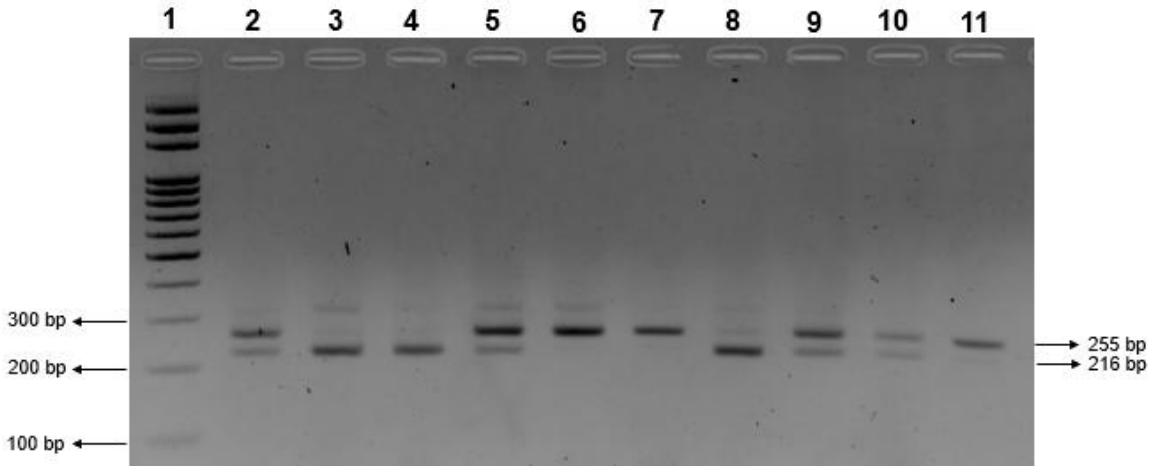


Figure 3.3: A 2% agarose gel showing the restriction profile of the *HLA-G* rs1710 polymorphism PCR products digested with PstI

Lane 1: 100 bp DNA ladder; Lane 2, 5, 9 and 10: participants heterozygous for the rs1710 SNP (GC genotype; fragment sizes: 255, 216 and 39); Lane 3, 4 and 8: participants homozygous for the rs1710 C allele (CC genotype; fragment sizes: 216 and 39 bp); Lane 6 and 7: participants homozygous for the rs1710 G allele (GG genotype; fragment sizes: 255 bp); Lane 11: undigested PCR amplicon (255 bp). The 39 bp fragment was not visible as it

was small and running in the dye front. The spurious band see above the 300bp mark did not influence genotyping.

3.3.2 Detection of the *HLA-G* 14 bp insertion/deletion polymorphism

Participants were genotyped for the presence of the *HLA-G* 14 bp indel by conventional PCR. Figure 3.4 illustrates the results obtained from the PCR amplification of the *HLA-G* gene 3' UTR containing the 14 bp insertion/deletion polymorphism.

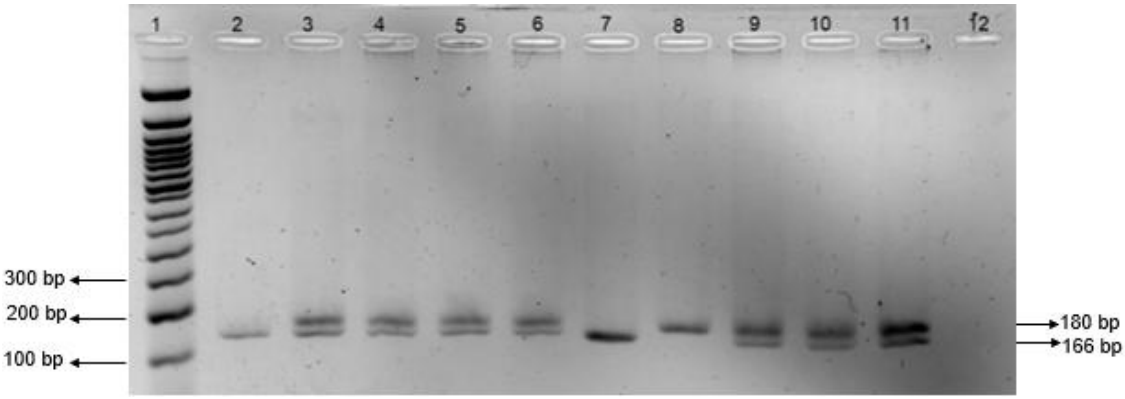


Figure 3.4: A 3% agarose gel image of PCR products obtained from the amplification of the *HLA-G* gene 3' UTR containing the 14 bp indel

Lane 1: 100 bp DNA ladder; Lane 2 and 7: participants homozygous for the deletion allele (DD genotype; fragment size of 166 bp); Lane 3-6 and 9-10: participants heterozygous for the insertion/deletion (ID genotype; fragment size 180 and 166 bp); Lane 8: participant homozygous for the insertion allele (II genotype; fragment size of 180 bp); Lane 12: NTC.

3.3.3 Detection of the *HLA-G* rs9380142 SNP

The rs9380142 A>G SNP was amplified using TaqMan probes specific to each allele. The real-time PCR amplification plot (Figure 3.5) was used to generate an allelic discrimination plot (Figure 3.6) from which genotypes were assigned.

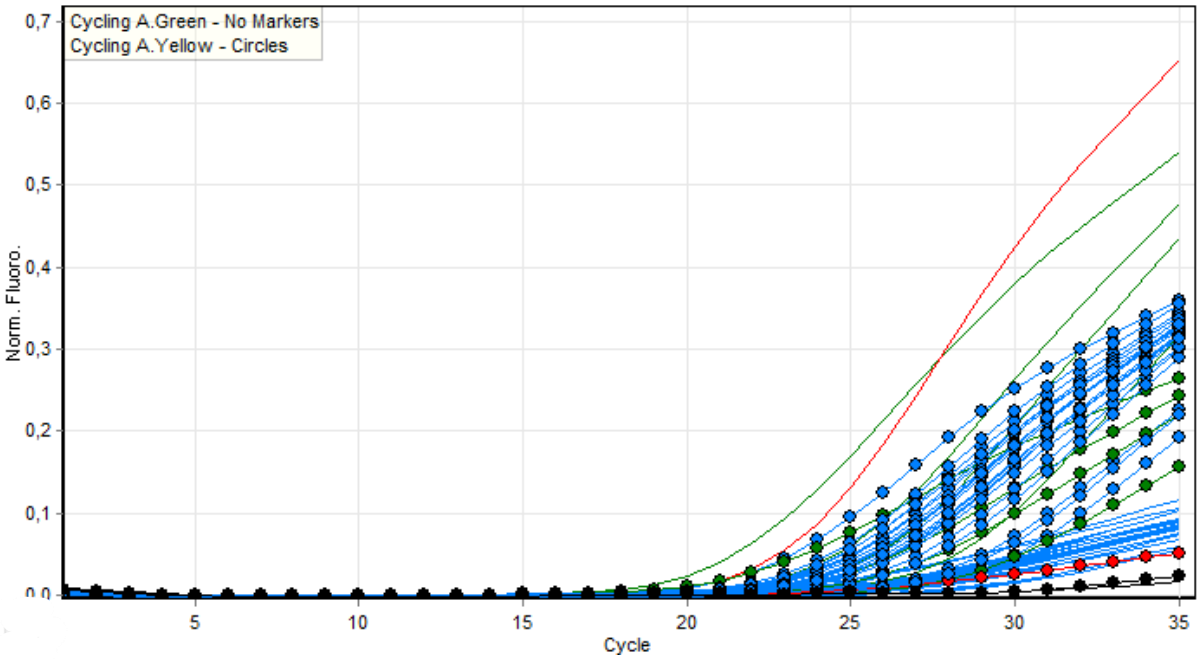


Figure 3.5: Amplification plot for the *HLA-G* rs9380142 SNP.

Each curve represents the amplification of the target sequence from a different sample based on the amount of fluorescence detected throughout 35 cycles. **Red:** amplification of the homozygous GG genotype. **Green:** amplification the homozygous AA genotype. **Blue:** amplification of the heterozygous AG genotype. **Black**: lack of amplification of the NTC.

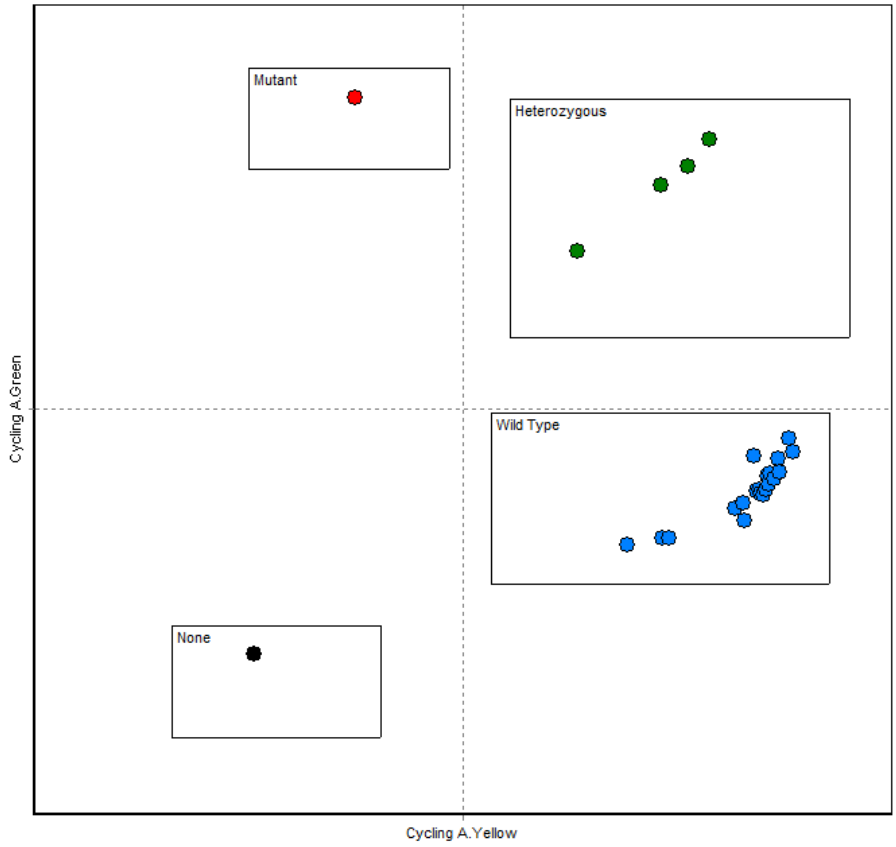


Figure 3.6: Allelic discrimination plot for the *HLA-G* rs9380142 SNP

Samples where fluorescence was detected only in the FAM™ channel (G allele) were grouped in the upper left quadrant and genotyped as GG; samples where fluorescence was detected only in the VIC® channel (A allele) were grouped in the bottom right quadrant; samples where fluorescence was detected in both the FAM™ and VIC® channels (AG heterozygote) were grouped in the upper right quadrant; samples (including the NTC) that showed no fluorescence were grouped in the bottom left quadrant.

3.3.4 Sequencing

Three samples representing the three possible genotypes for the rs1063320, rs1710 and 14 bp indel polymorphisms were selected and sent for Sanger sequencing. The sequencing results confirmed the genotyping results for the three polymorphisms (Figure 3.7 - 3.9). The rs9380142 assay was validated by the company and therefore samples were not sequenced for this polymorphism.

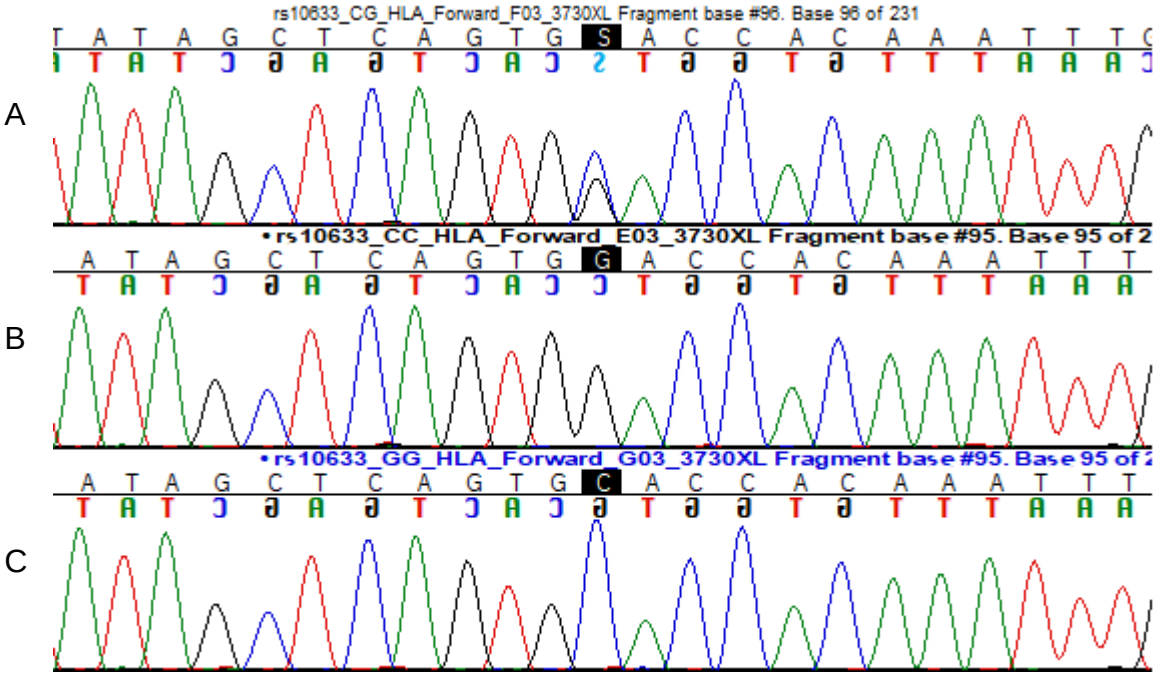


Figure 3.7: Chromatograms obtained from the amplification of the *HLA-G* gene containing the rs1063320 polymorphism. This SNP was sequenced in the forward and reverse direction.

A: chromatogram illustrating a heterozygous CG individual for the rs1063320 SNP

B: chromatogram illustrating a homozygous GG individual for the rs1063320 SNP

C: chromatogram illustrating a homozygous CC individual for the rs1063320 SNP

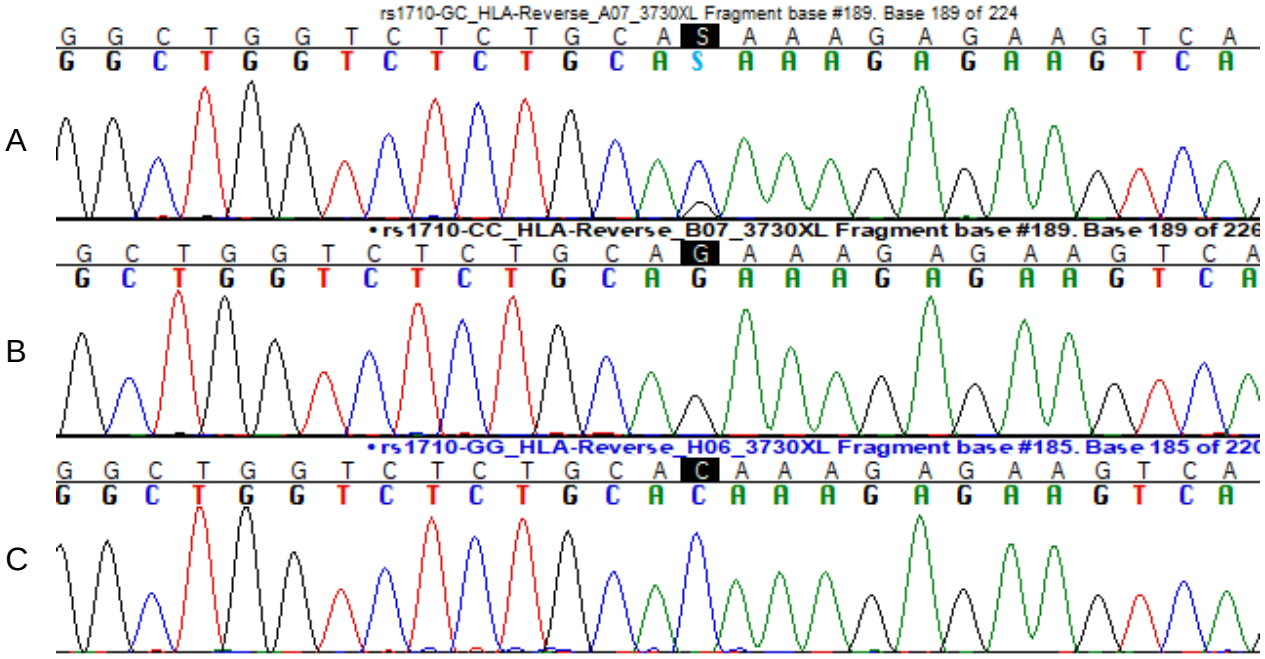


Figure 3.8: Chromatograms obtained from the amplification of the *HLA-G* gene containing the rs1710 polymorphism. This SNP was sequenced in the forward and reverse direction.

A: chromatogram illustrating a heterozygous CG individual for the rs1710 SNP

B: chromatogram illustrating a homozygous GG individual for the rs1710 SNP

C: chromatogram illustrating a homozygous CC individual for the rs1710 SNP

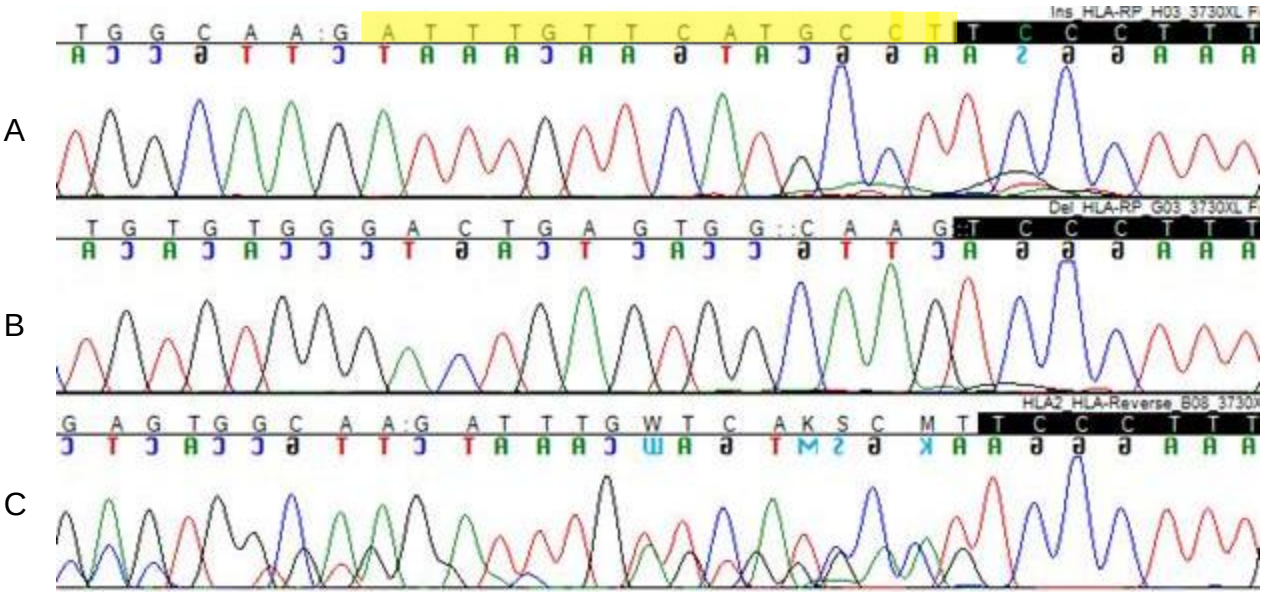


Figure 3.9: Chromatograms obtained from the amplification of the *HLA-G* gene containing the 14 bp indel polymorphism

A: chromatogram illustrating an individual homozygous for the 14 bp insertion

B: chromatogram illustrating an individual homozygous for the 14 bp indel

C: chromatogram illustrating an individual heterozygous for the 14 bp deletion

3.4 Genotypic and allelic frequencies of the *HLA-G* gene polymorphisms

The rs1063320, rs1710 and the 14 bp indel polymorphisms were in HWE (p=0.354, p=0.439 and p=0.457, respectively (Appendix E). However, the rs9380142 SNP was not in HWE (p=0.007) but for the purpose of this dissertation, further analyses were performed.

In the total study cohort, the G allele was the major allele for rs1063320, and the C allele was the major allele for rs1710 with a frequency of 0.61 and 0.60, respectively. The major allele for the 14 bp indel was the deletion allele with a frequency of 0.58 and the A allele was the major allele for rs9380142 with a frequency of 0.61. There was no significant difference in the genotypic and allelic frequencies between T1D cases and controls for any of the *HLA-G* polymorphisms (Table 3.3). Thus, these polymorphisms are not associated with an increased risk of developing T1D in the black South African population.

Table 3.3: Genotypic and allelic frequencies of the *HLA-G* gene polymorphisms in participants with T1D and controls

Polymorphism	Frequency (n)		p value
	Cases	Controls	
rs1063320			
Genotype			
CC	0.18 (46)	0.15 (30)	0.758
CG	0.45 (119)	0.46 (93)	
GG	0.37 (98)	0.39 (78)	
Allele			
C	0.40 (211)	0.38 (153)	0.525
G	0.60 (315)	0.62 (249)	
rs1710			
Genotype			
GG	0.18 (47)	0.15 (31)	0.699
GC	0.45 (118)	0.48 (97)	
CC	0.37 (98)	0.36 (73)	
Allele			
G	0.40 (212)	0.40 (159)	0.817
C	0.60 (314)	0.60 (243)	
14 bp indel			
Genotype			
II	0.18 (47)	0.18 (37)	0.251
ID	0.44 (116)	0.51 (102)	
DD	0.38 (100)	0.31 (62)	
Allele			
I	0.40 (210)	0.44 (176)	0.238
D	0.60 (316)	0.56 (226)	
rs9380142			
Genotype			
AA	0.52 (26)	0.51 (24)	0.610

AG	0.46 (23)	0.49 (23)	
GG	0.02 (1)	0.00 (0)	
Allele			
A	0.75 (75)	0.76 (71)	0.932
G	0.25 (25)	0.25 (23)	

3.5 The association of *HLA-G* gene polymorphisms with clinicopathological variables in the whole population

We determined if the genotypes of the *HLA-G* rs1063320, rs1710, 14 bp indel and rs9380142 polymorphisms were associated with any clinicopathological variables in the whole study population (Table 3.4 - 3.7). The deletion/deletion genotype of the 14 bp indel was associated with a significantly higher SBP in black South Africans. No other significant differences were found between clinicopathological variables and the four polymorphisms.

Table 3.4: Association of *HLA-G* rs1063320 genotypes with clinicopathological variables

Variable	**rs1063320 genotypes**			**p value**
	CC (n=76)	**CG (n=212)**	**GG (n=176)**	
Sex				
Male (%, n)	47.4 (36)	43.9 (93)	43.2 (76)	0.822
Female (%, n)	52.6 (40)	56.1 (119)	56.8 (100)	
BMI (kg/m^2)	25.9 ± 6.83	26.2 ± 6.75	25.7 ± 6.23	0.772
Glucose (mmol/L)	6.30 [4.60; 8.40]	6.50 [4.80; 9.20]	6.00 [4.70; 9.50]	0.640
SBP (mmHg)	128.6 ± 17.9	126 ± 16.1	126.7 ± 18.5	0.659
DBP (mmHg)	78.1 ± 12.0	77.8 ± 11.3	78.7 ± 12.0	0.853
HLA-G conc (ng/ml)	39.0 [21.8; 72.7]	40.6 [26.0; 63.9]	40.5 [26.4; 66.1]	0.757

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

Table 3.5: Association of *HLA-G* rs1710 genotypes with clinicopathological variables

Variable	rs1710 genotypes			p value
	GG (n=78)	GC (n=215)	CC (n=171)	
Sex				
Male (%, n)	47.4 (37)	44.2 (95)	42.7 (73)	0.783
Female (%, n)	52.6 (41)	55.8 (120)	57.3 (98)	
BMI (kg/m^2)	25.8 ± 6.76	26.2 ± 6.74	25.7 ± 6.27	0.736
Glucose (mmol/L)	6.10 [4.70; 8.40]	6.40 [4.80; 9.20]	6.10 [4.65; 9.80]	0.665
SBP (mmHg)	129.2 ± 18.2	125.8 ± 16.1	126.5 ± 18.4	0.508
DBP (mmHg)	78.4 ± 12.1	77.8 ± 11.3	78.6 ± 11.9	0.855
HLA-G conc (ng/ml)	39.0 [21.9; 71.8]	39.5 [25.2; 61.8]	41.6 [27.9; 72.4]	0.491

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

Table 3.6: Association of *HLA-G* 14 bp indel genotypes with clinicopathological variables

Variable	14 bp genotypes			p value
	II (n=84)	ID (n=218)	DD (n=162)	
Sex				
Male (%, n)	44 (37)	42.2 (92)	46.9 (76)	0.658
Female (%, n)	56 (47)	57.8 (126)	53.1 (86)	
BMI (kg/m^2)	25.6 ± 6.30	26.1 ± 6.69	25.9 ± 6.55	0.870
Glucose (mmol/L)	6.70 [4.75; 9.85]	6.00 [4.60; 8.80]	6.35 [4.80; 9.80]	0.781
SBP (mmHg)	124.4 ± 16.8	124.9 ± 17.1	130.1 ± 17.4	0.042
DBP (mmHg)	76.4 ± 12.0	78.3 ± 12.1	78.9 ± 10.9	0.473
HLA-G conc (ng/ml)	44.5 [25.6; 66.6]	39.4 [26.2; 65.3]	41.2 [25.2; 71.0]	0.545

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

Table 3.7: Association of *HLA-G* rs9380142 genotypes with clinicopathological variables

Variable	rs9380142 genotypes		p value
	AA (n=50)	AG+GG (n=47)	
Sex			
Male (%, n)	40.0 (20)	42.6 (20)	0.799
Female (%, n)	60.0 (30)	57.4 (27)	
BMI (kg/m^2)	25.3 ± 7.27	23.6 ± 5.98	0.214
Glucose (mmol/L)	5.40 [4.70; 8.70]	5.00 [4.10; 7.20]	0.529
SBP (mmHg)	120.4 ± 12.1	123.2 ± 14.8	0.448
DBP (mmHg)	74.1 ± 8.98	77.4 ± 9.58	0.190
HLA-G conc (ng/ml)	51.0 [34.4; 82.2]	35.3 [28.0; 55.2]	0.115

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

3.6 The association of *HLA-G* gene polymorphisms with clinicopathological variables in participants with T1D

We investigated the association of the four *HLA-G* polymorphisms with clinicopathological variables in cases (Table 3.8 - 3.11). No significant differences were seen between any of the clinicopathological variables and the rs1063320, rs1710, 14 bp indel and rs9380142 polymorphisms. Similarly, no association between *HLA-G* genotypes and any clinicopathological variables were seen in the control population (data not shown).

Table 3.8: Association of the *HLA-G* rs1063320 genotypes with clinicopathological variables in participants with T1D

Variable	rs1063320 genotypes			p value
	CC (n=46)	CG (n=119)	GG (n=98)	
Age at diagnosis (years)	18.7 ± 9.54	19.5 ± 9.81	19.4 ± 10.8	0.898
Duration of disease (years)	8.00 [3.00; 12.0]	5.00 [2.00; 12.0]	6.00 [2.00; 12.0]	0.644
Sex				
Male (%, n)	39.1 (18)	50.4 (60)	43.9 (43)	0.370
Female (%, n)	60.9 (28)	49.6 (59)	56.1 (55)	
BMI (kg/m^2)	24.8 ± 7.50	24.8 ± 6.37	25.5 ± 6.94	0.707
Glucose (mmol/L)	7.80 [4.60; 10.4]	8.25 [5.10; 14.5]	9.30 [5.30; 13.5]	0.622
HbA1C (%)	9.89 ± 2.31	10.8 ± 2.84	9.98 ± 2.90	0.067
SBP (mmHg)	129.1 ± 18.6	126.9 ± 16.6	128.0 ± 19.3	0.769
DBP (mmHg)	78.5 ± 12.7	77.9 ± 11.3	78.7 ± 12.4	0.902
HLA-G conc (ng/ml)	61.3 [20.8; 98.1]	52.8 [33.9; 94.5]	48.2 [34.4; 85.7]	0.778

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, HbA1c = glycated haemoglobin, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

Table 3.9: Association of the *HLA-G* rs1710 genotypes with clinicopathological variables in participants with T1D

Variable	rs1710 genotypes			p value
	GG (n=47)	GC (n=118)	CC (n=98)	
Age at diagnosis (years)	18.6 ± 9.43	19.7 ± 9.93	19.2 ± 10.8	0.831
Duration of disease (years)	8.00 [3.00; 12.0]	5.00 [2.00; 11.5]	6.50 [2.00; 12.0]	0.561
Sex				
Male (%, n)	38.3 (18)	51.7 (61)	42.9 (42)	0.217
Female (%, n)	61.7 (29)	48.3 (57)	57.1 (56)	
BMI (kg/m^2)	24.8 ± 7.42	24.8 ± 6.42	25.5 ± 6.91	0.723
Glucose (mmol/L)	7.80 [4.60; 10.4]	8.25 [5.15; 15.1]	9.30 [5.30; 13.0]	0.638
HbA1C (%)	9.82 ± 2.33	10.8 ± 2.83	10.0 ± 2.90	0.058
SBP (mmHg)	129.8 ± 18.9	127.0 ± 16.5	127.5 ± 19.3	0.678
DBP (mmHg)	78.8 ± 12.8	78.0 ± 11.3	78.4 ± 12.4	0.920
HLA-G conc (ng/ml)	60.2 [21.8; 94.0]	52.8 [32.6; 88.9]	50.2 [34.4; 88.4]	0.955

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, HbA1c = glycated haemoglobin, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

Table 3.10: Association of the *HLA-G* 14 bp indel genotypes with clinicopathological variables in participants with T1D

Variable	14 bp genotypes			p value
	II (n=47)	ID (n=116)	DD (n=100)	
Age at diagnosis (years)	17.5 ± 9.77	19.6 ± 10.4	19.8 ± 9.96	0.390
Duration of disease (years)	6.00 [3.00; 11.0]	5.00 [2.00; 12.0]	6.00 [2.00; 12.0]	0.766
Sex				
Male (%, n)	42.6 (20)	48.3 (59)	42.0 (42)	0.373
Female (%, n)	57.4 (27)	51.7 (57)	58.0 (58)	
BMI (kg/m^2)	25.3 ± 7.09	24.9 ± 6.38	25.1 ± 7.12	0.950
Glucose (mmol/L)	8.00 [4.80; 14.1]	8.65 [5.30; 14.0]	8.55 [5.40; 11.8]	0.870
HbA1C (%)	10.3 ± 2.65	10.6 ± 3.10	10.0 ± 2.47	0.319
SBP (mmHg)	124.3 ± 17.0	126.0 ± 18.3	131.2 ± 17.6	0.051
DBP (mmHg)	76.3 ± 11.9	77.9 ± 12.6	79.7 ± 11.1	0.287
HLA-G conc (ng/ml)	51.4 [30.8; 94.6]	52.3 [33.9; 84.3]	50.2 [30.8; 88.4]	0.875

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, HbA1c = glycated haemoglobin, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

Table 3.11: Association of the *HLA-G* rs9380142 genotypes with clinicopathological variables in participants with T1D

Variable	rs9380142 genotypes		p value
	AA (n=26)	AG+GG (n=24)	
Age at diagnosis (years)	18.3 ± 9.80	15.2 ± 10.4	0.286
Duration of disease (years)	3.50 [1.00; 15.0]	6.00 [2.00; 11.0]	0.254
Sex			
Male (%, n)	46.2 (12)	45.8 (11)	0.982
Female (%, n)	53.8 (14)	54.2 (13)	
BMI (kg/m^2)	23.6 ± 6.82	22.5 ± 5.81	0.520
Glucose (mmol/L)	8.00 [5.15; 11.0]	6.70 [3.90; 15.6]	0.962
HbA1C (%)	11.1 ± 3.87	10.6 ± 2.73	0.673
SBP (mmHg)	122.0 ± 12.2	125.7 ± 15.2	0.401
DBP (mmHg)	74.6 ± 9.17	77.0 ± 8.35	0.413
HLA-G conc (ng/ml)	46.9 [35.0; 70.2]	88.1 [37.5; 107.9]	0.325

Parametric data is represented as mean ± standard deviation and as median [lower and upper quartiles] for non-parametric data. BMI = body mass index, HbA1c = glycated haemoglobin, SBP = systolic blood pressure, DBP = diastolic blood pressure, *HLA-G*= human leukocyte antigen-G

3.7 The association of *HLA-G* gene polymorphisms in participants with T1D with an age at diagnosis before and after 21 years of age

To determine whether the alleles and/or the genotypes of the *HLA-G* rs1063320, rs1710, 14 bp indel and rs9380142 polymorphisms were associated with the age at diagnosis of T1D, cases were classified as either having an age at diagnosis <21years or ≥21 years (Table 3.12). The age at diagnosis categories were selected based on the threshold used by Padoa *et al* (2020) in a South African cohort with T1D. There was no association with any of the *HLA-G* genotypes investigated and age at diagnosis category in black South African participants with T1D.

Table 3.12: Comparison of genotypic and allelic frequencies of *HLA-G* polymorphisms between participants with T1D and an early and late age at diagnosis

Polymorphisms	Age at diagnosis frequency (n)		p value
	<21 years	≥ 21 years	
rs1063320			
Genotype			
CC	0.19 (28)	0.16 (18)	0.703
CG	0.43 (63)	0.47 (54)	
GG	0.38 (56)	0.37 (42)	
Allele			
C	0.40 (119)	0.39 (90)	0.817
G	0.60 (175)	0.61 (138)	
rs1710			
Genotype			
GG	0.20 (29)	0.16 (18)	0.510
GC	0..41 (61)	0.48 (55)	
CC	0.39 (57)	0.36 (41)	
Allele			
G	0.40 (119)	0.40 (91)	0.896
C	0.60 (175)	0.60 (137)	
14 bp indel			
Genotype			
DD	0.37 (54)	0.39 (45)	0.099
ID	0.41 (60)	0.48 (55)	
II	0.22 (33)	0.12 (14)	
Allele			
D	0.57 (168)	0..64 (145)	0.136
I	0.43 (126)	0.36 (83)	
rs9380142			
Genotype			
AA	0.45 (14)	0.63 (12)	0.216
AG + GG	0.55 (17)	0.37 (7)	

Allele			
A	0.71 (44)	0.82 (31)	0.234
G	0.29 (18)	0.18 (7)	

3.8 Logistic regression model determining variables associated with T1D

A backward, stepwise logistic regression analysis was performed to ascertain the variables associated with T1D. When controlling for sex, BMI, rs1063320, rs1710 and 14 bp indel polymorphisms, there was a positive association between serum HLA-G concentration and T1D (p value <0.001). Participants with T1D are 5.2 times more likely to have higher HLA-G concentrations compared to controls. The rs9380142 SNP was excluded from the model due to small sample size.

Table 3.13: Logistic regression analysis for the determinants of T1D status

Dependent variable	**Independent variable**	**Odds ratio (95% CI)**	**p value**
T1D status[a]	LogHLA-G concentration	5.15 (2.37 – 11.2)	<0.001

For full model: p<0.001 *(n =304),* [a]*control = 0, case = 1*

3.9 Logistic regression model determining variables associated with age at diagnosis category

The determinants of age at diagnosis of T1D in our study population were investigated in a logistic regression analysis. After controlling for sex, rs1063320 and rs1710 polymorphisms, the insertion/insertion genotype was associated with an earlier age at diagnosis of T1D (p =0.042).

Table 3.14: Logistic regression analysis for the determinants of age at diagnosis of T1D

Dependent variable	Independent variable	Odds ratio (95% CI)	p value
Age at diagnosis[a]	HLA-G 14 bp indel polymorphism[b]	0.90 (0.81-1.00)	0.042

For full model p < 0.031 (n = 261), [a]*age at diagnosis <21 years = 0, age at onset ≥ 21 years = 1;* [b]*homozygous insertion genotype = 0, homozygous deletion and heterozygous indel genotypes = 1*

3.10 Multivariable linear regression model to determine factors affecting serum HLA-G concentrations

A multivariable linear regression analysis was performed to ascertain the main determinants of SBP (Table 3.15). T1D status, age, BMI and the 14 bp indel were associated with SBP when controlling for sex and rs1063320 and rs1710 polymorphisms.

Participants with T1D had a SBP that was approximately 5.62 mmHg higher than that of controls. For every one-year increase in age and every unit (kg/m^2) increase in BMI, the SBP increased by 0.412 mmHg and 0.67 mmHg, respectively. Participants with the 14 bp deletion/deletion genotype had SBP readings 4.00 mmHg higher than participants with the homozygous insertion and heterozygous 14 bp indel genotypes.

Table 3.15: Multivariable linear regression analysis for determinants of systolic blood pressure

Dependent variable	Independent variable	B value	p value
SBP	T1D status[a]	5.620	0.030
	Age	0.412	<0.001
	BMI	0.667	<0.001
	14 bp indel polymorphism[b]	-4.004	0.036

For the full model R^2 = 0.075, p< 0.001 (n= 304); [a]control = 0, case = 1, [b]homozygous deletion genotype = 0, homozygous insertion and heterozygous indel genotypes = 1

4. Discussion

Type 1 diabetes mellitus (T1D) is an autoimmune disease that is characterized by a T cell mediated destruction of insulin secreting pancreatic β-cells (Amodio *et al*., 2021). The cause of T1D is multifactorial but has a major genetic component. The *HLA-G* gene was identified as an independent locus for disease susceptibility (Eike *et al.*, 2009). The HLA-G molecule is constitutively expressed in the pancreas, and it plays an important role in the suppression of immune responses (Arnaiz-Villena *et al*., 2021). Polymorphisms in the *HLA-G* gene, depending on location, affect transcription factor binding, peptide binding, isoform production, mRNA expression and consequently HLA-Gs immunoregulatory functions (Castelli *et al*., 2014; Donadi *et al*., 2011) suggesting a role in T1D pathogenesis and/or disease progression.

Sequencing of the *HLA-G* gene in cases and controls identified 103 variants, 33 of which were novel variants. The *HLA-G* rs1063320, rs1710, 14 bp indel and rs9380142 polymorphisms were not associated with T1D in the black South African population. However, higher HLA-G concentrations were associated with T1D in the black South African population. The 14 bp homozygous insertion genotype was associated with an earlier age at diagnosis. In addition, the deletion/deletion genotype of the 14 bp indel, BMI, age and T1D status were associated with SBP.

4.1 Participant characteristics

The blood glucose concentrations were higher in cases than controls. Hyperglycaemia is a known consequence of T1D as it is a result of the T cell mediated destruction of pancreatic β-cells that leads to an absolute insulin deficiency (Amodio *et al*., 2021). Insulin is a peptide hormone that stimulates the utilization and storage of glucose by promoting glucose uptake from circulation into skeletal muscle cells and adipose tissue via the glucose transporter, GLUT4 (Satoh, 2014). Moreover, within the liver, insulin activates glycogen synthesis and decreases gluconeogenic gene expression (Petersen and Shulman, 2018). Therefore, an insulin deficiency results in decreased cellular glucose uptake and utilization, increased hepatic gluconeogenesis (production of glucose from non-carbohydrate sources) and glycogenolysis (degradation of glycogen into glucose), which all consequently lead to the development of hyperglycaemia (Satoh, 2014).

In addition, cases had high HbA1c results (10.3 ± 2.80 %) confirming poor glycaemic control within the three months prior to sampling. Poor glycaemic control is defined by an HbA1c ≥ 8 % (Mata-Cases *et al*., 2020). However, as these participants are being managed in the clinic, and are on insulin treatment, better glycaemic control would be expected.

The controls had significantly higher BMIs than cases. An older age and female sex have previously been associated with adiposity (Mancuso and Bouchard, 2019). However, in our study, cases and controls were matched for age (p=0.823) and sex (p=0.398). Thus, these variables cannot account for the differences seen in BMI. The lower BMI in cases may be explained by their poor glycaemic control, which will lead to reduced storage and greater breakdown of triglycerides in adipose tissues (Dilworth *et al*., 2021) due to the increased activity of lipase, a hormone that is normally suppressed by insulin (Satoh, 2014).

The SBP was significantly higher in cases than controls (p=0.021). This significance remained after controlling for confounding variables in multivariate analysis and is discussed in section 4.8.

The HLA-G serum concentrations were significantly higher in cases than controls (p <0.001). This finding was confirmed in multivariate analysis and will be discussed further in section 4.5.

4.2 Identification of variants in the *HLA-G* gene of black South African participants

Sequencing of the *HLA-G* gene identified a total of 103 variants and of these, 70 were identified as known variants and 33 as novel. The majority (56%) of known variants were found in exons (15 synonymous, 23 missense, and 1 nonsense mutation). The remainder of the known variants were found in introns (37%) and the 3' UTR (7%). For the unknown variants, 52% were found in exons (7 synonymous and 11 missense mutations) and 48% in intronic regions. Variants that occur in intronic regions may be involved in alternative splicing which can give rise to different isoforms that exhibit different properties (Donadi *et al*., 2011). Variants in the promoter region and 3' UTR can influence miRNA binding thus affecting mRNA expression. Variants that occur in

the coding region of a gene can affect the biological function of the protein by altering its binding affinity for its receptor or changing the 3-dimensional structure of the protein which could prevent dimerization (Castelli *et al*., 2021).

Across the human genome, the *HLA* complex is known as the most polymorphic region (Tshabalala *et al*., 2022). The *HLA-G* gene coding region is relatively conserved in comparison to the classic *HLA* genes (Xu *et al*., 2020). Studies have shown that the most variation occurs in the regulatory regions and intronic regions (de Albuquerque *et al.*, 2016; Castelli *et al*., 2021; Chen *et al*., 2023). In our study we did not sequence the promoter region, or the majority of the intronic regions, therefore important variants may have been missed. In addition, black South Africans and the African population as a whole, exhibit greater genetic diversity than other populations due to various factors such as migrations, genetic admixture, and local evolutionary pressures (Tshabalala *et al*., 2022; Mellet *et al*., 2019). This may explain the high number of novel variants seen in our small group of participants selected for sequencing.

4.3 Allelic and genotypic frequencies of the four *HLA-G* polymorphisms

All the polymorphisms, except rs9380142, were in HWE. Therefore, our cohort's allelic and genotypic frequencies for rs9380142 are not considered to be representative of the whole population. It is likely that rs9380142 was not in HWE due to the small number of participants genotyped for this polymorphism (n= 97).

The frequencies of the major alleles in our study for rs1063320, rs1710 and the 14 bp indel were similar to that of the African population from the 1000 Genomes Project (Table 7.2, Appendix E) (The 1000 Genomes Project Consortium, 2015). However, the frequency of rs9380142 was significantly different to the 1000 genome African population data. This deviation in expected frequencies may be explained by the SNP not being in HWE and the small sample size.

4.4 The *HLA-G* polymorphisms are not associated with T1D in the black South African population

None of the HLA-G polymorphisms investigated in the current study were associated with T1D in black South Africans.

Similarly, studies conducted in the South Indian (Gunavathy *et al*., 2023) and Brazilian (de Albuquerque *et al*., 2016) populations found that the rs1063320 SNP was not associated with T1D. The rs1063320 SNP occurs at position +3142 in the 3' UTR of the *HLA-G* gene and results in a C to G nucleotide change (Castelli *et al*., 2009). This SNP is believed to alter the binding affinity of specific miRNAs to the *HLA-G* 3' UTR, thus influencing *HLA-G* gene expression levels (Castelli *et al*., 2009). The presence of the G allele increases the affinity of hsa-miR-148a, hsa-miR-148b, and hsa-miR-152 miRNAs to the 3' UTR of the *HLA-G* mRNA, downregulating the expression of *HLA-G* (Castelli *et al*., 2009) and thus increasing the risk of T1D.

In contrast to our findings, a study conducted in a Brazilian cohort found that the rs1710 CC genotype was significantly decreased in participants with T1D compared to controls and was associated with a 15.5 % decreased risk of developing T1D (de Albuquerque *et al.*, 2016). The *HLA-G* rs1710 SNP occurs at position +3010 in the 3' UTR of the *HLA-G* gene and results in a G to C nucleotide change (Castelli *et al*., 2009). The rs1710 C allele has been associated with alternative splicing which generates a 92 bp deletion from the start of exon 8 resulting in a truncated mRNA that is more stable (Castelli *et al*., 2023). Thus, in the presence of the rs1710 C allele, there is an increase in HLA-G levels (Castelli *et al*., 2023) which is protective against T1D due to its immunosuppressive effect on T cells.

A study conducted in a Cypriot population, supporting our findings, found no significant difference in the frequency of the 14 bp indel polymorphism between participants with T1D and controls (Gerasimou *et al*., 2016). In addition, a study by de Albuquerque *et al.* (2016) reported that there was no association between the 14 bp indel and T1D in the Brazilian population. However, a second Brazilian study found the deletion allele and the homozygous deletion genotype to be associated with T1D (Silva *et al.,* 2016). A familial study among the South Indian population, also showed an association of the 14 bp deletion allele and deletion/deletion genotype with an increased risk of T1D especially in females (Gunavathy *et al*., 2023). This may have been due to gender specific factors such as progesterone which have been reported to regulate the expression of *HLA-G* through the presence of a progesterone response element at the 5′ upstream region of the gene (Yie *et al*., 2006). In contrast, the homozygous insertion

genotype was reported to confer a 3.82-fold increased risk of developing T1D in the Iranian population (Rezaei *et al*., 2019).

To the best of our knowledge, there is only one study on the rs9380142 SNP and T1D (de Albuquerque *et al.*, 2016) and similar to our study, they found no association between rs9380142 and T1D in the Brazilian population.

The different associations seen in the literature may be due to the genetic heterogeneity /ethnic backgrounds of the study participants or the sample size investigated.

4.5 Serum HLA-G levels are associated with T1D in the black South African population

The HLA-G molecule is constitutively expressed in insulin secretory granules and is upregulated on the cell surface of islet cells that have been stimulated to secrete insulin (Cirulli *et al*., 2006). HLA-G inhibits T lymphocyte and NK cell differentiation, proliferation and cytotoxicity, cytokine secretion and immunoglobulin production upon binding to their specific inhibitory receptors (Amiot *et al*., 2014). In our study we found an association between higher serum HLA-G concentrations and T1D status. These results are unexpected as high HLA-G levels are thought to be protective of T1D development due to the immune-tolerogenic role of HLA-G.

There are a limited number of studies investigating HLA-G levels in T1D with conflicting findings. Wyatt and colleagues (2019) reported that in pancreatic tissue from participants with recent-onset T1D, an upregulation of *HLA-G* at protein level was seen in the islets, in both α and β-cells. In contrast, another study compared the expression of *HLA-G* between histologically normal pancreatic (HNP) tissues and pancreatic tissue from participants with T1D and found significantly higher HLA-G staining in the islets from HNP samples compared to those from T1D participants (Bertol *et al*., 2019).

In addition, HLA-G molecules must form homodimers to exert their immunosuppressive functions (Donadi *et al*., 2011). HLA-G acts as a ligand for inhibitory receptors (e.g. immunoglobulin-g-like transcript 2 (ILT2)) that are expressed

on various immune cells such as dendritic, NK and T cells (Shiroshi *et al*., 2003). The dimerization of HLA-G increases exposure to the ILT2 binding site by altering the structure, resulting in increased HLA-G/ILT binding affinity to $CD8^+$ T cells (Attia *et al*., 2020). This results in increased expression and secretion of Fas ligand, inducing cell apoptosis in $CD8^+$ T and NK cells and inhibiting the activity of cytotoxic T cells protecting against β-cell destruction (Contini *et al*., 2003). The high levels of HLA-G in our cohort are not protective against T1D. Therefore, it is possible the HLA-G dimers are not forming thus HLA-G is not able to exert its immune-tolerogenic effects resulting in β-cell destruction and progression of T1D. Alternatively, as the participants with T1D in our cohort had hyperglycaemia which is known to result in inflammation (Butkowski *et al*., 201), the higher HLA-G levels seen in cases compared to controls may be as a result of the ongoing inflammatory response. The association of hyperglycaemia and HLA-G levels is supported by a study that looked at the effect of glucose concentrations on *HLA-G* expression. They found that the higher the glucose concentration, the higher the HLA-G concentration (Solini *et al*., 2010).

4.6 The 14 bp insertion is associated with an earlier age at diagnosis of T1D in the black South African population

In a univariate analysis none of the four polymorphisms studied were associated with age at diagnosis, however, on multivariate analysis the 14 bp insertion/insertion genotype was associated with an earlier age at diagnosis. It has been shown in literature that the presence of the 14 bp insertion results in decreased levels of HLA-G (Rousseau *et al*., 2003, Hviid *et al*.,2004, 2006; Chen *et al*., 2008). Low HLA-G concentrations affect its tolerogenic role. It is therefore hypothesised that low HLA-G levels increase an individual's risk of an earlier age at diagnosis of T1D as $CD8^+$ cytotoxic cells and NK cells are no longer inhibited, resulting in more rapid β-cell destruction. However, this does not explain our results as there was no association between the 14 bp indel polymorphism and HLA-G concentrations. In contrast to our finding, a Cypriot study found the homozygous deletion genotype to be associated with an earlier age of onset amongst the participants with T1D (Geramisou *et al*., 2016). The frequency of the homozygous deletion genotype was almost threefold higher in early onset participants with T1D (0-8 years of age) compared to those with a late onset (13-39 years of age). Similarly, the homozygous deletion genotype was

associated with earlier age of onset in a Brazilian population (Silva *et al*., 2016). In the South African black population, it has been found that the age at onset falls into two peaks, the earlier peak at 11-15 years and a later peak at 26-30 years (Padoa *et al*., 2020). Therefore, we do not see many participants in the 0-8 years category, and it is possible that this is why our results are inconsistent with the Silva study. Alternatively, the discrepancy between our results and those published in the literature may be explained by the14 bp indel polymorphism being in linkage disequilibrium with the true causative gene/polymorphism. It is therefore possible that the earlier age at onset seen in the studies may be due to HLA-G being unable to provide sufficient immune-tolerance in the pancreas (Geramisou *et al*., 2016).

There was a lack of association of *HLA-G* rs9380142, rs1063320 and rs1710 SNPs with age at diagnosis. To the best of our knowledge, no other study has investigated these three polymorphisms in the context of age at diagnosis.

4.7 The 14 bp indel polymorphism is associated with SBP in the South African black population.

Black South Africans with the 14 bp homozygous deletion genotype had significantly higher SBP than those with the insertion allele. To the best of our knowledge, no other study has investigated the association of this SNP with SBP in the context of T1D. However, a Mexican study found that among their participants with type 2 diabetes mellitus, those with the 14 bp insertion/insertion genotype were almost twice as likely to present with high blood pressure than those with the homozygous deletion (García-González *et al*., 2014).

Since there is no extensive research on this relationship, the mechanism is unknown. We hypothesise that the 14 bp indel polymorphism affects blood pressure via the VEGF/endothelin-1 pathway. The 14 bp deletion mRNA transcript is known to be expressed at significantly higher levels than the insertion transcript (Rousseau *et al*., 2003). In addition, it has been shown that the soluble HLA-G1 isoform inhibits vascular endothelial growth factor (VEGF) mediated angiogenesis and endothelial cell proliferation through apoptosis (Fons *et al*., 2006). VEGF is known to promote nitric oxide production thus, inhibition of VEGF signalling leads to a decrease in nitric oxide

(endothelial dysfunction) and activation of the endothelin-1 pathways which result in vasoconstriction and raised blood pressure (de Jesus-Gonzalez *et al*., 2012).

4.8 Age, BMI and T1D status are associated with SBP in the South African black population

In our study cohort, for every one-year increase in age, SBP increased by 0.412 mmHg. Our results are in agreement with data from the Framingham Heart Study which found that after the age of 30 years, SBP increased with age (Franklin and Wong, 2013). This finding is expected as ageing has been associated with: 1) arterial stiffness due to loss of elastin and deposition of collagen fibres (Vatner *et al*., 2021), 2) endothelial dysfunction characterised by a loss of nitric oxide production and hence decreased vasodilation (Higashi *et al*., 2012) and 3) gradual loss of smooth muscle function due to the decreased expression of a principal receptor for nitric oxide, soluble guanylyl cyclase, which contributes to the development of hypertension (Thijssen *et al*., 2016).

For every 1 unit (kg/m^2) increase in BMI, the SBP increased by 0.67 mmHg. This finding is supported by a study conducted in South-Eastern Poland that found SBP to be significantly associated with BMI (Leszczak *et al*., 2024). Similarly, Shukuri *et al*. (2019) found that individuals with a BMI > 25 kg/m^2 were 4.3 times more likely to be hypertensive compared to participants with a BMI between 18-25 kg/m^2. Obesity is associated with endothelial dysfunction and one of the mechanisms for this relationship is the secretion of proinflammatory cytokines by perivascular adipose tissue which leads to a reduction in nitric oxide availability resulting in an increase in vascular resistance and a raised blood pressure (Mengozzi *et al*., 2020).

Participants with T1D had a SBP that was approximately 5.62 mmHg higher than that of controls. It is well known that individuals living with diabetes have a 1.5 to 2-fold increased risk of developing hypertension than the general population (Tatsumi and Ohkubo, 2017). In individuals with T1D, the development of microalbuminuria usually precedes the development of hypertension, thereafter the hypertension drives diabetic nephropathy, which in turn increases BP (Mogensen, 2003; Jia and Sowers, 2021). Hypertension is also associated with T1D in individuals with a normal albumin

excretion rate where isolated systolic hypertension is three times more common than in controls (Rönnback *et al*., 2004). This could be attributed to the increased arterial stiffness caused by the effect of hyperglycaemia on the vascular system (Cheung and Li, 2012) such as the deposition of advanced glycation end products and cross linking of collagen molecules in the arterial wall (Rönnback *et al*., 2004). Moreover, it has been suggested that there is accelerated vascular aging in individuals with T1D, as age related changes i.e., increase in systolic BP and decrease in diastolic BP were noted to occur earlier (Rönnback *et al*., 2004).

4.9 Limitations of the study

One of the limitations to our study was that sequencing of the *HLA-G* gene exons was carried out on a relatively small sample, and this may have prevented the detection of other variants that may be associated with T1D and/or age at diagnosis in the black South African population. In addition, as we did not sequence the promoter and all intronic regions of the *HLA-G* gene, we may have missed polymorphisms in these regions that are associated with T1D. The rs9380142 polymorphism was not in HWE possibly due to the small sample size. Moreover, this may have prevented the detection of associations of the polymorphism with T1D and/or age at diagnosis.

This study did not look at whether the polymorphisms under investigation were in linkage disequilibrium with each other or with any other polymorphisms within the *HLA-G* gene. In addition, the effects of haplotypes on the risk of development of T1D were not analysed.

5. Conclusions

In conclusion, our study revealed that the *HLA-G* rs1063320, rs1710, 14 bp indel and rs9380142 polymorphisms are not associated with T1D in the black South African population. The presence of the 14 bp insertion/insertion genotype was however, associated with an earlier age at diagnosis. These results are supportive of the involvement of *HLA-G* in the earlier age at diagnosis of T1D. In addition, the deletion/deletion genotype of the 14 bp indel was associated with a significantly higher SBP in black South Africans potentially due to the inhibition of angiogenesis and endothelial cell proliferation.

There was no association of *HLA-G* genotypes with serum HLA-G concentrations. However, the cases were 5.2 times more likely to have higher HLA-G concentrations compared to controls in the black South African population. The exact mechanism behind this association is not known prompting the need for future studies.

The association of HLA-G rs1063320, rs1710, rs9380142 and 14 bp indel polymorphisms with T1D is inconsistent among various ethnic populations. Future studies looking at these associations should be conducted in a larger cohort to confirm our findings. Future studies should include haplotype analysis and expression studies in a larger cohort.

6. References

Abdul-Hussein, S.S., Ali, E.N., Alkhalidi, N.M., Zaki, N.H. and Ad'hiah, A.H., 2020. Susceptibility role of soluble HLA-G and *HLA-G* 14-bp insertion/deletion polymorphism in inflammatory bowel disease. *Egyptian journal of medical human genetics*, *21*(1): 1-12.

Alegre, E., Rizzo, R., Bortolotti, D., Fernandez-Landázuri, S., Fainardi, E. and González, A., 2014. Some basic aspects of HLA-G biology. *Journal of immunology research*, 2014.

Amiot, L., Vu, N. and Samson, M., 2014. Immunomodulatory properties of HLA-G in infectious diseases. *Journal of immunology research*, *2014*.

Amodio, G., Mandelli, A., Curto, R., Rancoita, P.M., Stabilini, A., Bonfanti, R., De Pellegrin, M., Bosi, E., Di Serio, C., Battaglia, M. and Gregori, S., 2021. Altered Frequency and Phenotype of HLA-G-Expressing DC-10 in Type 1 Diabetes Patients at Onset and in Subjects at Risk to Develop the Disease. *Frontiers in immunology*, 12(750162).

American Diabetes Association, 2023. Standards of care in diabetes—2023 abridged for primary care providers. *Clinical Diabetes*, *41*(1): 4-31.

Arnaiz-Villena, A., Juarez, I., Suarez-Trujillo, F., López-Nares, A., Vaquero, C., Palacio-Gruber, J. and Martin-Villa, J.M., 2021. HLA-G: Function, polymorphisms and pathology. *International journal of immunogenetics*, *48*(2): 172-192.

Atkinson, M.A., 2012. The pathogenesis and natural history of type 1 diabetes. *Cold Spring Harbor perspectives in medicine*, *2*(11): a007641.

Atkinson, M.A., Eisenbarth, G.S. and Michels, A.W., 2014. Type 1 diabetes. *The lancet*, *383*(9911): 69-82.

Attia, J.V., Dessens, C.E., van de Water, R., Houvast, R.D., Kuppen, P.J. and Krijgsman, D., 2020. The molecular and functional characteristics of HLA-G and the interaction with its receptors: where to intervene for cancer immunotherapy?. *International Journal of Molecular Sciences*, 21(22): 8678.

Balcha, S.A., Demisse, A.G., Mishra, R., Vartak, T., Cousminer, D.L., Hodge, K.M., Voight, B.F., Lorenz, K., Schwartz, S., Jerram, S.T. and Gamper, A., 2020. Type 1 diabetes in Africa: an immunogenetic study in the Amhara of North-West Ethiopia. *Diabetologia*, *63*(10): 2158-2168.

Baschal, E.E., Sarkar, S.A., Boyle, T.A., Siebert, J.C., Jasinski, J.M., Grabek, K.R., Armstrong, T.K., Babu, S.R., Fain, P.R., Steck, A.K. and Rewers, M.J., 2011. Replication and further characterization of a Type 1 diabetes-associated locus at the telomeric end of the major histocompatibility complex. *Journal of diabetes*, *3*(3): 238-247.

Bertol, B.C., Dias, F.C., da Silva, D.M., Ramalho, L.N.Z. and Donadi, E.A., 2019. Human Antigen Leucocyte (HLA)-G and HLA-E are differentially expressed in pancreatic disorders. *Human Immunology*, *80*(11): 948-954.

Butkowski, E.G. and Jelinek, H.F., 2017. Hyperglycaemia, oxidative stress and inflammatory markers. *Redox Report*, *22*(6), pp.257-264.

Carosella, E.D., Moreau, P., Le Maoult, J., Le Discorde, M., Dausset, J. and Rouas-Freiss, N., 2003. HLA-G molecules: from maternal-fetal tolerance to tissue acceptance. *Advances in immunology*, 81: 199-252.

Castelli, E.C., de Almeida, B.S., Muniz, Y.C., Silva, N.S., Passos, M.R., Souza, A.S., Page, A.E., Dyble, M., Smith, D., Aguileta, G. and Bertranpetit, J., 2021. HLA-G genetic diversity and evolutive aspects in worldwide populations. *Scientific reports*, *11*(1): 23070.

Castelli, E.C., Moreau, P., e Chiromatzo, A.O., Mendes-Junior, C.T., Veiga-Castelli, L.C., Yaghi, L., Giuliatti, S., Carosella, E.D. and Donadi, E.A., 2009. In silico analysis of microRNAS targeting the HLA-G 3′ untranslated region alleles and haplotypes. *Human immunology*, 70(12): 1020-1025.

Castelli, E.C., Paes, G.S., da Silva, I.M., Moreau, P. and Donadi, E.A., 2023. The+ 3010/C single nucleotide polymorphism (rs1710) at the HLA-G 3′ untranslated region is associated with a short transcript exhibiting a deletion of 92 nucleotides. *Immunogenetics*, *75*(2): 155-160.

Castelli, E.C., Veiga-Castelli, L.C., Yaghi, L., Moreau, P. and Donadi, E.A., 2014. Transcriptional and posttranscriptional regulations of the HLA-G gene. *Journal of immunology research*, 2014.

Cirulli, V., Zalatan, J., McMaster, M., Prinsen, R., Salomon, D.R., Ricordi, C., Torbett, B.E., Meda, P. and Crisa, L., 2006. The class I HLA repertoire of pancreatic islets comprises the nonclassical class Ib antigen *HLA-G*. *Diabetes*, *55*(5): 1214-1222.

Chen, X.Y., Yan, W.H., Lin, A., Xu, H.H., Zhang, J.G. and Wang, X.X., 2008. The 14 bp deletion polymorphisms in HLA-G gene play an important role in the expression of soluble HLA-G in plasma. *Tissue antigens*, *72*(4): 335-341.

Chen, D.P., Wang, P.N., Hour, A.L., Lin, W.T., Hsu, F.P., Wang, W.T. and Tseng, C.P., 2023. The association between genetic variants at 3'-UTR and 5'-URR of HLA-G gene and the clinical outcomes of patients with leukemia receiving hematopoietic stem cell transplantation. *Frontiers in Immunology*, *14*: 1093514.

Cheung, B.M. and Li, C., 2012. Diabetes and hypertension: is there a common metabolic pathway?. *Current atherosclerosis reports*, 14: 160-166.

Contini, P., Ghio, M., Poggi, A., Filaci, G., Indiveri, F., Ferrone, S. and Puppo, F., 2003. Soluble HLA-A,-B,-C and-G molecules induce apoptosis in T and NK $CD8^+$ cells and inhibit cytotoxic T cell activity through CD8 ligation. *European journal of immunology*, *33*(1): 125-134.

Dahl, M., Klitkou, L., Christiansen, O.B., Djurisic, S., Piosik, Z.M., Skovbo, P., Møller, A.M., Steffensen, R. and Hviid, T.V.F., 2015. Human leukocyte antigen (HLA)-G during pregnancy part II: associations between maternal and fetal HLA-G genotypes and soluble HLA-G. *Human immunology*, 76(4): 260-271.

de Albuquerque, R.S., Mendes-Junior, C.T., Lucena-Silva, N., da Silva, C.L.L., Rassi, D.M., Veiga-Castelli, L.C., Foss-Freitas, M.C., Foss, M.C., Deghaide, N.H.S., Moreau, P. and Gregori, S., 2016. Association of HLA-G 3′ untranslated region variants with type 1 diabetes mellitus. *Human immunology*, *77*(4): 358-364.

de Jesus-Gonzalez, N., Robinson, E., Moslehi, J. and Humphreys, B.D., 2012. Management of antiangiogenic therapy-induced hypertension. *Hypertension*, *60*(3): 607-615.

Dilworth, L., Facey, A. and Omoruyi, F., 2021. Diabetes mellitus and its metabolic complications: the role of adipose tissues. *International journal of molecular sciences*, *22*(14): 7644.

Donadi, E.A., Castelli, E.C., Arnaiz-Villena, A., Roger, M., Rey, D. and Moreau, P., 2011. Implications of the polymorphism of *HLA-G* on its function, regulation, evolution and disease association. *Cellular and molecular life sciences*, *68*(3): 369-395.

Eike, M.C., Olsson, M., Undlien, D.E., Dahl-Jørgensen, K., Joner, G., Rønningen, K.S., Thorsby, E. and Lie, B.A., 2009. Genetic variants of the HLA-A, HLA-B and AIF1

loci show independent associations with type 1 diabetes in Norwegian families. *Genes and immunity*, 10(2): 141-150.

Fons, P., Chabot, S., Cartwright, J.E., Lenfant, F., L'Faqihi, F., Giustiniani, J., Herault, J.P., Gueguen, G., Bono, F., Savi, P. and Aguerre-Girr, M., 2006. Soluble HLA-G1 inhibits angiogenesis through an apoptotic pathway and by direct binding to CD160 receptor expressed by endothelial cells. *Blood*, *108*(8): 2608-2615.

Franklin, S.S. and Wong, N.D., 2013. Hypertension and cardiovascular disease: contributions of the Framingham Heart Study. *Global heart*, *8*(1): 49-57.

Galicia-Garcia, U., Benito-Vicente, A., Jebari, S., Larrea-Sebal, A., Siddiqi, H., Uribe, K.B., Ostolaza, H. and Martín, C., 2020. Pathophysiology of type 2 diabetes mellitus. *International journal of molecular sciences*, 21(17): 6275.

García-González, I.J., Valle, Y., Rivas, F., Figuera-Villanueva, L.E., Muñoz-Valle, J.F., Flores-Salinas, H.E., Gutiérrez-Amavizca, B.E., Dávalos-Rodríguez, N.O. and Padilla-Gutiérrez, J.R., 2014. The 14 bp Del/Ins HLA-G polymorphism is related with high blood pressure in acute coronary syndrome and type 2 diabetes mellitus. *BioMed Research International*, 2014.

Gerasimou, P., Skordis, N., Picolos, M., Spyridonidis, A. and Costeas, P., 2016. HLA-G 14-bp polymorphism affects the age of onset in Type I Diabetes Mellitus. *International journal of immunogenetics*, 43(3): 135-142.

Gunavathy, N., Asirvatham, A., Chitra, A. and Jayalakshmi, M., 2023. Evaluation of HLA-G 14 bp Ins/Del and+ 3142 C/G Polymorphisms in Type 1 Diabetes among South Indian Population. *Indian Journal of Endocrinology and Metabolism*, *27*(3): 223-229.

Harrison, G.A., Humphrey, K.E., Jakobsen, I.B. and Cooper, D.W., 1993. A 14 bp deletion polymorphism in the *HLA-G* gene. *Human molecular genetics*, *2*(12): 2200-2200.

Higashi, Y., Kihara, Y. and Noma, K., 2012. Endothelial dysfunction and hypertension in aging. *Hypertension Research*, *35*(11), pp.1039-1047.

Hviid, T.V.F., Rizzo, R., Christiansen, O.B., Melchiorri, L., Lindhard, A. and Baricordi, O.R., 2004. HLA-G and IL-10 in serum in relation to HLA-G genotype and polymorphisms. *Immunogenetics*, *56*: 135-141.

Jia, G. and Sowers, J.R., 2021. Hypertension in diabetes: an update of basic mechanisms and clinical disease. *Hypertension*, *78*(5): 1197-1205.

Jothikumar, N., Kang, G. and Hill, V.R., 2009. Broadly reactive TaqMan® assay for real-time RT-PCR detection of rotavirus in clinical and environmental samples. *Journal of virological methods*, *155*(2): 126-131.

Krijgsman, D., Roelands, J., Hendrickx, W., Bedognetti, D. and Kuppen, P.J., 2020. HLA-G: a new immune checkpoint in cancer?. *International journal of molecular sciences*, 21(12): 4528.

Kongsbak, M., Levring, T.B., Geisler, C. and Von Essen, M.R., 2013. The vitamin D receptor and T cell function. *Frontiers in immunology*, 4(148).

Kukreja, A. and Maclaren, N.K., 1999. Autoimmunity and diabetes. *The Journal of Clinical Endocrinology & Metabolism*, *84*(12): 4371-4378.

Leszczak, J., Czenczek-Lewandowska, E., Asif, M., Baran, J., Mazur, A. and Wyszyńska, J., 2024. Risk factors and prevalence of hypertension in older adults from south-eastern Poland: an observational study. *Scientific Reports*, *14*(1): 1450.

Mancuso, P. and Bouchard, B., 2019. The impact of aging on adipose function and adipokine synthesis. *Frontiers in endocrinology*, 10: 437303.

Martín-Villa, J.M., Vaquero-Yuste, C., Molina-Alejandre, M., Juarez, I., Suárez-Trujillo, F., López-Nares, A., Palacio-Gruber, J., Barrera-Gutiérrez, L., Fernández-Cruz, E., Rodríguez-Sainz, C. and Arnaiz-Villena, A., 2022. HLA-G: too much or too little? Role in cancer and autoimmune disease. *Frontiers in immunology*, *13*, p.796054.

Mashal, S., Khanfar, M., Al-Khalayfa, S., Srour, L., Mustafa, L., Hakooz, N.M., Zayed, A.A., Khader, Y.S. and Azab, B., 2021. SLC30A8 gene polymorphism rs13266634 associated with increased risk for developing type 2 diabetes mellitus in Jordanian population. *Gene*, 768:145279.

Mellet, J., Tshabalala, M., Agbedare, O., Meyer, P.W.A., Gray, C.M. and Pepper, M.S., 2019. Human leukocyte antigen (HLA) diversity and clinical applications in South Africa. *South African Medical Journal*, *109*(8 Supplement 1): S30-S34.

Mengozzi, A., Masi, S. and Virdis, A., 2020. Obesity-related endothelial dysfunction: Moving from classical to emerging mechanisms. *Endocrine and Metabolic Science*, *1*(3-4): 100063.

Mogensen, C.E., 2003. Microalbuminuria and hypertension with focus on type 1 and type 2 diabetes. *Journal of internal medicine*, *254*(1): 45-66.

Petersen, M.C. and Shulman, G.I., 2018. Mechanisms of insulin action and insulin resistance. *Physiological reviews*, 98(4):2133-2223.

Padoa, C.J., Rheeder, P., Pirie, F.J., Motala, A.A., Van Dyk, J.C. and Crowther, N.J., 2020. Identification of a subgroup of black South Africans with type 1 diabetes who are older at diagnosis but have lower levels of glutamic acid decarboxylase and islet antigen 2 autoantibodies. *Diabetic medicine*, *37*(12): 2067-2074.

Redondo, M.J., Steck, A.K. and Pugliese, A., 2018. Genetics of type 1 diabetes. *Pediatric diabetes*, *19*(3):346-353.

Rewers, M. and Ludvigsson, J., 2016. Environmental risk factors for type 1 diabetes. *The lancet*, *387*(10035): 2340-2348.

Rezaei, F., Zareei, N., Razmi, N., Nikeghbalian, S. and Azarpira, N., 2019. Genetic Polymorphism of *HLA-G* 14-bp Insertion/Deletion in Pancreas Transplant Recipients and Its Association With Type 1 Diabetes Mellitus. *Experimental and clinical transplantation: Official journal of the middle east society for organ transplantation,* 19(2): 154-159.

Rizzo, R., Bortolotti, D., Bolzani, S. and Fainardi, E., 2014. HLA-G molecules in autoimmune diseases and infections. *Frontiers in immunology*, *5*: 592.

Rodrigues, J.K.F., Crovella, S. and da Silva, R.C., 2020. The *HLA-G* 14 bp allele frequency in different populations: A global meta-analysis. *Meta gene*, 23(100624).

Rönnback, M., Fagerudd, J., Forsblom, C., Pettersson-Fernholm, K., Reunanen, A. and Groop, P.H., 2004. Altered age-related blood pressure pattern in type 1 diabetes. *Circulation*, *110*(9): 1076-1082.

Rousseau, P., Le Discorde, M., Mouillot, G., Marcou, C., Carosella, E.D. and Moreau, P., 2003. The 14 bp deletion-insertion polymorphism in the 3′ UT region of the HLA-G gene influences HLA-G mRNA stability. *Human immunology*, *64*(11): 1005-1010.

Satoh, T., 2014. Molecular mechanisms for the regulation of insulin-stimulated glucose uptake by small guanosine triphosphatases in skeletal muscle and adipocytes. *International journal of molecular sciences*, *15*(10), pp.18677-18692.

SEMDSA Type 2 Diabetes Guidelines Expert Committee, 2017. SEMDSA 2017 guidelines for the management of type 2 diabetes mellitus. *Journal of endocrinology, metabolism and diabetes of South Africa*, 22(1 Suppl 1): S1-S196.

Scarabel, L., Garziera, M., Fortuna, S., Asaro, F., Toffoli, G. and Geremia, S., 2020. Soluble HLA-G expression levels and HLA-G/irinotecan association in metastatic colorectal cancer treated with irinotecan-based strategy. *Scientific reports*, 10(1): 1-11.

Silva, H.P.V., Ururahy, M.A.G., Souza, K.S.C., Loureiro, M.B., Oliveira, Y.M.C., Oliveira, G.H.M., Luchessi, A.D., Carvalho, K.T.C., Freitas, J.C.O.C., Donadi, E.A. and Hirata, R.D.C., 2016. The association between the *HLA-G* 14-bp insertion/deletion polymorphism and type 1 diabetes. *Genes & immunity*, *17*(1): 13-18.

Solini, A., Muscelli, E., Stignani, M., Melchiorri, L., Santini, E., Rossi, C., Astiarraga, B.D., Rizzo, R. and Baricordi, O.R., 2010. Soluble human leukocyte antigen-g expression and glucose tolerance in subjects with different degrees of adiposity. *The Journal of Clinical Endocrinology & Metabolism*, *95*(7): 3342-3346.

Solis-Herrera, C., Triplitt, C., Reasner, C., DeFronzo, R.A. and Cersosimo, E., 2018. Classification of diabetes mellitus. *Endotext [Internet]*. https://www.ncbi.nlm.nih.gov/books/NBK279119/

Shukuri, A., Tewelde, T. and Shaweno, T., 2019. Prevalence of old age hypertension and associated factors among older adults in rural Ethiopia. *Integrated blood pressure control*, 23-31.

Sun, H., Saeedi, P., Karuranga, S., Pinkepank, M., Ogurtsova, K., Duncan, B.B., Stein, C., Basit, A., Chan, J.C., Mbanya, J.C. and Pavkov, M.E., 2022. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes research and clinical practice*, 183:109119.

Tatsumi, Y. and Ohkubo, T., 2017. Hypertension with diabetes mellitus: significance from an epidemiological perspective for Japanese. *Hypertension Research*, *40*(9): 795-806.

The 1000 Genomes Project Consortium, 2015.A global reference for human genetic variation. *Nature*, 526(7571): 68-74.

Thijssen, D.H., Carter, S.E. and Green, D.J., 2016. Arterial structure and function in vascular ageing: are you as old as your arteries?. *The Journal of physiology*, *594*(8), pp.2275-2284.

Tshabalala, M., Mellet, J., Vather, K., Christoffels, A. and Pepper, M.S., 2022. High resolution HLA~ a,~ B,~ C,~ DRB1,~ DQA1, and~ DQB1 diversity in south African populations. *Frontiers in Genetics*, *13*: 711944.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., 2012. Primer3—new capabilities and interfaces. *Nucleic acids research*, *40*(15): e115-e115.

van de Water, R.B., Krijgsman, D., Houvast, R.D., Vahrmeijer, A.L. and Kuppen, P.J., 2021. A critical assessment of the association between HLA-G expression by

carcinomas and clinical outcome. *International Journal of Molecular Sciences*, *22*(15), p.8265.

Vatner, S.F., Zhang, J., Vyzas, C., Mishra, K., Graham, R.M. and Vatner, D.E., 2021. Vascular stiffness in aging and disease. *Frontiers in Physiology*, *12*, p.762437.

Verkauskiene, R., Danyte, E., Dobrovolskiene, R., Stankute, I., Simoniene, D., Razanskaite-Virbickiene, D., Seibokaite, A., Urbonaite, B., Jurgeviciene, N., Vitkauskiene, A. and Schwitzgebel, V., 2016. The course of diabetes in children, adolescents and young adults: does the autoimmunity status matter?. *BMC endocrine disorders*, *16*(1): 1-13.

Wållberg, M. and Cooke, A., 2013. Immune mechanisms in type 1 diabetes. *Trends in immunology*, *34*(12), pp.583-591.

Wang, D.D. and Hu, F.B., 2018. Precision nutrition for prevention and management of type 2 diabetes. *The lancet diabetes & endocrinology*, 6(5): 416-426.

Wyatt, R.C., Lanzoni, G., Russell, M.A., Gerling, I. and Richardson, S.J., 2019. What the HLA-I—Classical and Non-classical HLA Class I and Their Potential Roles in Type 1 Diabetes. *Current diabetes reports*, *19*(12): 1-11.

Xu, H.H., Yan, W.H. and Lin, A., 2020. The Role of HLA-G in human papillomavirus infections and cervical carcinogenesis. *Frontiers in Immunology*, 11: 1349.

Yie, S.M., Xiao, R. and Librach, C.L., 2006. Progesterone regulates HLA-G gene expression through a novel progesterone response element. *Human reproduction*, *21*(10): 2538-2544.

7. Appendices

7.1 Appendix A: Stock solutions and dilution preparations for PCR

Primer working solution (10 µM)

10 µl Primer stock solution (100 µM)

90 µl dH2O

TBE Buffer (10x)

109 g Tris base

55.7 g Boric acid

7.45 g Ethylenediaminetetraacetic acid (EDTA)

Make up to 1 litre with distilled water (dH_2O)

TBE Buffer (1x)

100 ml 10x TBE

900 ml dH_2O

TE buffer (1x)

10 mM Tris-HCl

1 mM EDTA

pH 8.0, in deionised water

Agarose gel (1 %)

1g agarose 100 ml

1 x TBE

Microwave for 2 minutes, cool slightly and add 0.5 µl Ethidium bromide to gel.

7.2 Appendix B: Ethics certificates

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

p

R14/49 Dr C Padoa

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M200174

NAME: (Principal Investigator)	Dr C Padoa
DEPARTMENT:	School of Pathology Department of Chemical Pathology Medical School University
PROJECT TITLE:	The phenotypic and genotypic characterization of patients diagnosed with diabetes mellitus in South Africa
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	Renewal of M150885
SUPERVISOR:	Not applicable
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2020/02/27

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the 3rd Floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to submit details to the Committee. I **agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in January and will therefore reports and re-certification will be due early in the month of January each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature Date

PLEASE QUOTE THE CLEARANCE CERTIFICATE NUMBER IN ALL ENQUIRIES

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R49 Ms CR Dzimbanhete

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M220484

NAME: (Principal Investigator)	Ms CR Dzimbanhete
DEPARTMENT:	School of Pathology Department of Chemical Pathology Medical School University
PROJECT TITLE:	*The association of HLA-G with Type 1 diabetes in the South African Black population*
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M20/01/74
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office - <Nicoleen.Potgieter@wits.ac.za>
SUPERVISOR:	Drs C Padoa and E Cave
APPROVED BY:	CB Penny Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2022/05/12

This Clearance Certificate is valid for 5 years from the date of approval. An extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office secretariat on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. I agree to submit a yearly progress report. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **2022/04/01** and therefore reports and re-certification will be due in the month of **2022/04/01** each year. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator Date

7.3 Appendix C: *HLA-G* variants identified from sequencing analysis

The tables below represent the sequencing results of all 8 exons. For every novel variant detected, it is highlighted in red followed by a portion of the adjacent sequence.

EXON ONE				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1630223 G A	 0.556 (10) 0.444 (8)	 0.833 (15) 0.167 (3)	G:49%; A:51%	Synonymous
rs1630185 G A	 0.556 (10) 0.444 (8)	 0.833 (15) 0.167 (3)	G:49%; A:51%	Synonymous

INTRON 1				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs56388903 G A	 0.944 (17) 0.056 (1)	 0.824 (28) 0.176 (6)	G:89%; A:11%	Intron variant

rs6932888 G C	 0.722 (13) 0.278 (5)	 0.647 (22) 0.353 (12)	G:83%; C:17%	Intron variant
rs6932596 C T	 0.722 (13) 0.278 (5)	 0.618 (21) 0.382 (13)	C:83%; T:17%	Intron variant
rs1629329 T C	 0.528 (19) 0.472 (17)	 0.824 (14) 0.176 (3)	C:62%; T:38%	Intron variant
rs1628628 C T	 0.528 (19) 0.472 (17)	 0.824 (14) 0.176 (3)	C:49%; T:51%	Intron variant

EXON TWO				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1130355 G A	 0.579 (11) 0.421 (8)	 0.833 (15) 0.167 (3)	G:49%; A:51%	Synonymous

rs113311833 A T	 0.972 (35) 0.028 (1)	 0.813 (26) 0.187 (6)	A:100%	Missense (glutamic acid > aspartate)
ACACCAAGGCCC A T	 0.971 (33) 0.029 (1)	 0.786 (11) 0.214 (3)	No data	Missense (asparagine > isoleucine)

INTRON 2				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1376975606 A G C	 0.75 (18) 0.25 (6) 0	 0. 786 (11) 0.071 (1) 0.143 (2)	A:100%	Intron variant
rs1246735144 T C	 0.833 (20) 0.167 (4)	 0.786 (11) 0.214 (3)	No data	Intron variant
rs766585360 C T	 0.875 (21) 0.125 (3)	 0.857 (12) 0.143 (2)	C:100%	Intron variant

EXON THREE				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs565858069 G C	 0.929 (26) 0.071 (2)	 0.792 (19) 0.208 (5)	G:100%	Missense (aspartate > histidine)

EXON FOUR				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1285867439 T C	 0.769 (10) 0.231 (3)	 0.75 (15) 0.25 (5)	T: 100%	Missense (tyrosine > histidine)
rs878913602 G A	 0.562 (9) 0.438 (7)	 0.227 (5) 0.773 (17)	G:100%	Missense (glutamic acid > lysine)
rs879086343 T	 0.722 (13)	 0.643 (18)	T: 100%	Missense

C	0.278 (5)	0.357 (10)		(isoleucine > threonine)
rs1344510974 G A	 0.444 (4) 0.556 (5)	 0.607 (17) 0.393 (11)	G: 100%	Missense (valine > methionine)
rs1443101185 C T	 0.611 (11) 0.389 (7)	 0.643 (9) 0.357 (5)	C: 100%	Synonymous
rs140470441 T A	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	T: 100%	Synonymous
rs775674432 G A	 0.5 (9) 0.5 (9)	 0.607 (17) 0.393 (11)	G: 100%	Synonymous
rs1761083619 T G	 0.778 (14) 0.222 (4)	 0.792 (19) 0.208 (5)	No data	Missense (tryptophan > arginine)
rs1218025680 G C A	 0.444 (4) 0.556 (5) 0	 0.370 (10) 0.407 (11) 0.222 (6)	G: 100%	Missense G>A/C (glutamic acid > lysine/glutamine)

rs79264452 G A	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	G: 100%	Synonymous
rs76951509 G A	 1 (9) 0	 0.929 (13) 0.071 (1)	G:99%; A:1%	Synonymous
rs74547057 T C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	T: 99%; C: 1%	Synonymous
rs1043833437 C A	 0.444 (4) 0.556 (5)	 0.679 (19) 0.321 (9)	C: 100%	Missense (proline > leucine)
rs17875406 G A C	 0.389 (7) 0.444 (8) 0.167 (3)	 0.607 (17) 0.393 (11)	G: 82%; A: 18%	Synonymous
rs1212068397 A G	 0.944 (17) 0.056 (1)	 0.893 (25) 0.107 (3)	A: 100%	Missense (methionine > valine)

rs571857419 T C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	T: 100%	Missense (methionine > threonine)
rs537772641 G C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	G: 100%	Missense (methionine > isoleucine)
rs1438202915 C T	 1 (9) 0	 0.857 (24) 0.143 (4)	C: 100%	Synonymous
rs759332633 A C	 0.5 (9) 0.5 (9)	 0.643 (9) 0.357 (5)	A: 100%	Synonymous
rs774859148 A G	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	A: 100%	Missense (lysine > glutamic acid)
GGAGCTCGTG G C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	No data	Synonymous
GCAGGGGAT G T	 0.833 (15) 0.167 (3)	 0.964 (27) 0.036 (1)	No data	Missense (alanine > serine)

GAAGTGGGCAG G A	 0.778 (14) 0.222 (4)	 0.714 (20) 0.286 (8)	No data	Synonymous
TGGTGCCTT T G	 0.833 (15) 0.167 (3)	 0.821 (23) 0.179 (5)	No data	Missense (valine > glycine)
GTGCAGCA G A	 0.556 (10) 0.444 (8)	 0.643 (9) 0.357 (5)	No data	Missense (valine > methionine)
GCCCCTCATG G A C	 0.833 (15) 0.167 (3) 0	 0.857 (24) 0.036 (1) 0.107 (3)	No data	G>A: Synonymous G>C: Missense (glutamic acid > aspartic acid)
GATGCA G A	 0.722 (13) 0.278 (5)	 0.821 (23) 0.179 (5)	No data	Missense (arginine > lysine)

INTRON 4				
	Frequency (n)			Type of variant
Accession no. or position	Cases	Controls	1000 Genomes frequency	
rs763816713 A G	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	A: 100%	Intron variant
rs767930873 C G	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	C:100%	Intron variant
rs750833031 A G	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	A: 100%	Intron variant
rs757697673 G A	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	No data	Intron variant
rs1581668804 T G	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	T:100%	Intron variant
rs1761105699 G	 0.444 (4)	 0.643 (9)	No data	Intron variant

A	0.556 (5)	0.357 (5)		
rs1302651565 C T	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	C:100%	Intron variant
rs1761107123 A G	 0.5 (4) 0.5 (4)	 0.5 (4) 0.5 (4)	No data	Intron variant
rs533677430 G A	 0.5 (4) 0.5 (4)	 0.5 (4) 0.5 (4)	G:100%	Intron variant
rs761503504 A T	 0.5 (4) 0.5 (4)	 0.607 (17) 0.393 (11)	A: 100%	Intron variant
rs1761112876 C T	 0.875 (7) 0.125 (1)	 0.857 (12) 0.143 (2)	No data	Intron variant
GTCTGTTAGGG G C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	No data	Intron variant
GGAAAGCAGG G	 0.444 (4)	 0.643 (9)	No data	Intron variant

A	0.556 (5)	0.357 (5)		
TCTCTGAAGA T C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	No data	Intron variant
AAGACCTTTAAC A G	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	No data	Intron variant
ACCTTTAACAGG A C	 0.444 (4) 0.556 (5)	 0.643 (9) 0.357 (5)	No data	Intron variant
TAACAGGGTCGG T C	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Intron variant
AGGGTCGGTGG T C	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Intron variant
TGGGTGAGGGCT T G	 0.438 (7) 0.562 (9)	 0.536 (15) 0.464 (13)	No data	Intron variant
GTGAGGGCTGGG G	 0.5 (4)	 0.643 (9)	No data	Intron variant

C	0.5 (4)	0.357 (5)		
CTGGGGTCAGA C T	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Intron variant
AGACCCTCACCTT A G	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Intron variant
ACCCTCACCTTCA A C	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Intron variant

EXON FIVE				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs761148652 G A	 0.875 (14) 0.125 (2)	 0.857 (24) 0.143 (4)	G: 100%	Missense (lysine > asparagine)
rs766671384 A C	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	A:100%	Missense (glutamine > proline)

rs1761115068 T A	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Missense (leucine > glutamine)
rs573045964 T C	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	T: 100%	Synonymous
rs1049033 C T	 0.75 (12) 0.25 (4)	 0.75 (21) 0.25 (7)	C: 75%; T: 25%	Synonymous
rs1385852462 C T	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	C: 100%	Missense (alanine > valine)
rs1761121390 A C	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	No data	Missense (threonine > serine)
rs796748483 C T	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	C: 100%	Missense (threonine > isoleucine)
rs768731691 C T	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	C: 100%	Missense (alanine > valine)

rs747314068 G A	 0.438 (9) 0.562 (7)	 0.615 (8) 0.385 (5)	G: 100%	Missense (alanine > threonine)
rs771824819 T C	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	T: 100%	Synonymous
rs1242436454 G A	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	G: 100%	Missense (valine > methionine)
rs1190802521 G A T	 0.60 (12) 0.20 (4) 0.20 (4)	 0.615 (16) 0.038 (1) 0.346 (9)	G: 100%	Nonsense (tryptophan > stop codon)
GTCTTCCCTGCCC G A	 0.5 (4) 0.5 (4)	 0.577(15) 0.308 (8) 0.115 (3)	No data	Synonymous
TTGTCCTTGC T C	 0.625 (10) 0.375 (6)	 0.643 (18) 0.357 (10)	No data	Missense (valine > alanine)
TGCAGCTG			No data	Synonymous

T G	0.5 (4) 0.5 (4)	0.643 (9) 0.357 (5)		
AGCTGTAG A T	 0.5 (4) 0.5 (4)	 0.643 (9) 0.357 (5)	No data	Synonymous
TAGTCACTGGA T C	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	No data	Missense (valine > alanine)
AGTCACTGGAG A T	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	No data	Synonymous
TGCTGTGGAGAA T A	 0.875 (7) 0.125 (1)	 0.923 (12) 0.077 (1)	No data	Missense (valine > glutamic acid)
CTGTGGAGAA C A	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	No data	Missense (leucine > methionine)
AGAAGAGCTCAG G A	 0.5 (4) 0.5 (4)	 0.615 (8) 0.385 (5)	No data	Missense (lysine > arginine)

INTRON 5				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1761127540 G A	 0.429 (3) 0.571 (4)	 0.346 (9) 0.654 (17)	No data	Intron variant
rs1243330298 C A	 0.571 (4) 0.429 (3)	 0.615 (8) 0.385 (5)	C:100%	Intron variant
rs777740260 A G	 0.571 (4) 0.429 (3)	 0.615 (8) 0.385 (5)	A:100%	Intron variant
TGGGGTCTGAG T C	 0.571 (4) 0.429 (3)	 0.583 (7) 0.417 (5)	No data	Intron variant
AGTTTTCTTGT A G	 0.571 (4) 0.429 (3)	 0.583 (7) 0.417 (5)	No data	Intron variant
rs1761132614 G	 0.571 (4)	 0.636 (7)	No data	Intron variant

A	0.429 (3)	0.364 (4)		
CAAATTGAAAAG C G	0.776 (14) 0.224 (4)	0.818 (18) 0.182 (4)	No data	Intron variant
AATTGAAAAGG A G	0.833 (15) 0.167 (3)	0.818 (18) 0.182 (4)	No data	Intron variant

EXON SIX				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1761172520 A C	0.611 (11) 0.389 (7)	0.682 (15) 0.318 (7)	No data	Missense (Aspartic acid > glycine)
rs763482973 G A	0.5 (9) 0.5 (9)	0.792 (19) 0.208 (5)	G: 100%	3’ UTR variant

INTRON 7				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs915670 G A	 0.667 (12) 0.333 (6)	 0.769 (20) 0.231 (6)	G: 75%; A:25%	Intron variant
rs915669 G T	 0.556 (5) 0.444 (4)	 0.769 (20) 0.213 (6)	G: 38%; T: 62%	Intron variant
rs915668 G C	 0.667 (12) 0.333 (6)	 0.423 (11) 0.577 (15)	C: 38%; G: 62%	Intron variant

EXON EIGHT				
	Frequency (n)			
Accession no. or position	**Cases**	**Controls**	**1000 Genomes frequency**	**Type of variant**
rs1710 G C	 0.333 (1) 0.667 (2)	 0.75 (9) 0.25 (3)	G:38%; C:62%	3’ UTR variant

rs17179108 C T	 0.5 (3) 0.5 (3)	 1 (6)	C:88%; T:12%	3’ UTR variant
rs138249160 T C	 0.333 (1) 0.667 (2)	 1 (6)	T:100%	3’ UTR variant
rs1063320 C G	 0.333 (1) 0.667 (2)	 0.75 (9) 0.25 (3)	C:38%; G:62%	3’ UTR variant

7.4 Appendix D: ELISA standard curve

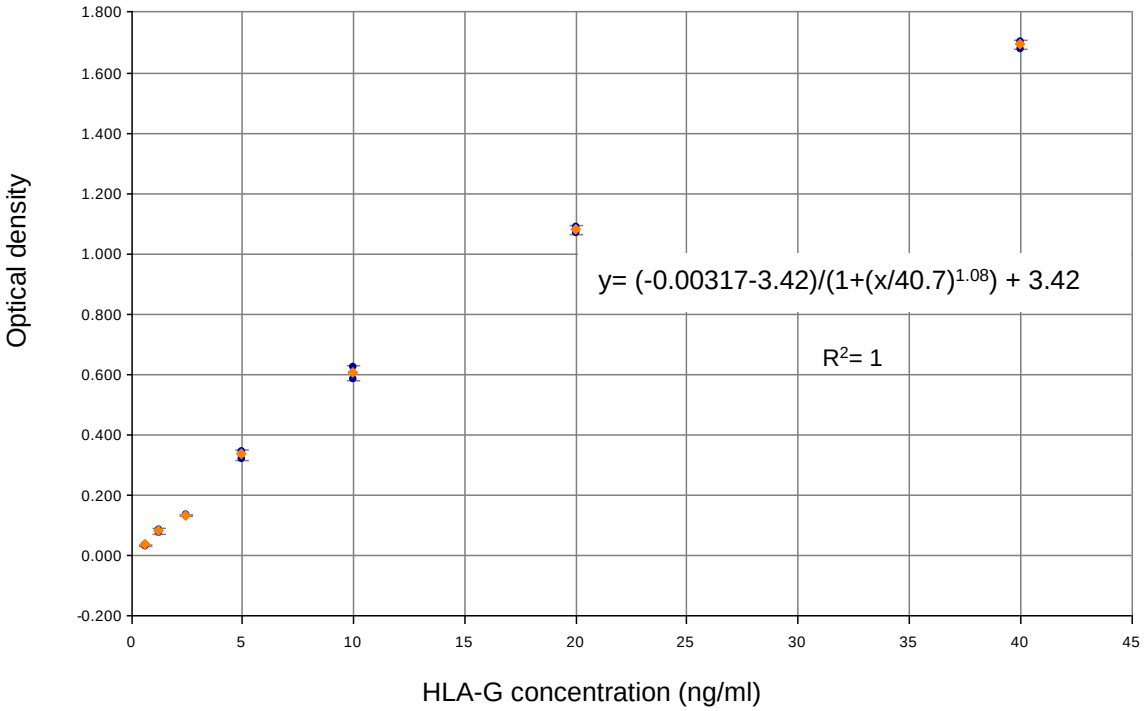


Figure 7.1: The four-parameter logistic (standard) curve used to calculate serum concentrations of HLA-G in different samples using OD determined from the ELISA assay

7.5 Appendix E: Hardy-Weinberg equilibrium

Table 7.1: Hardy-Weinberg equilibrium

Polymorphism	Participants' frequency (n)	Chi^2 (χ^2)	p value
rs1063320			
CC	0.163 (76)	0.861	0.354
CG	0.456 (212)		
GG	0.381 (177)		
rs1710			
GG	0.168 (78)	0.598	0.439
GC	0.462 (215)		
CC	0.370 (172)		
14 bp indel			
II	0.181 (84)	0.553	0.457
ID	0.469 (218)		
DD	0.350 (163)		
rs9380142			
AA	0.515 (50)	7.25	0.007*
AG	0.474 (46)		
GG	0.011 (1)		

Table 7.2: Comparison of HLA-G polymorphism frequencies between our cohort and the 1000 genome African population

Polymorphism	**Frequency**	
	Our study	**1000 Genome African**
rs1063320		
G allele	0.61	0.62
rs1710		
C allele	0.60	0.62
14 bp indel		
Deletion allele	0.58	0.62
rs9380142		
A allele	0.61	0.81*

*p<0.005 frequencies significantly different when compared tour study

7.6 Appendix F: PCR gel images of amplified exon regions

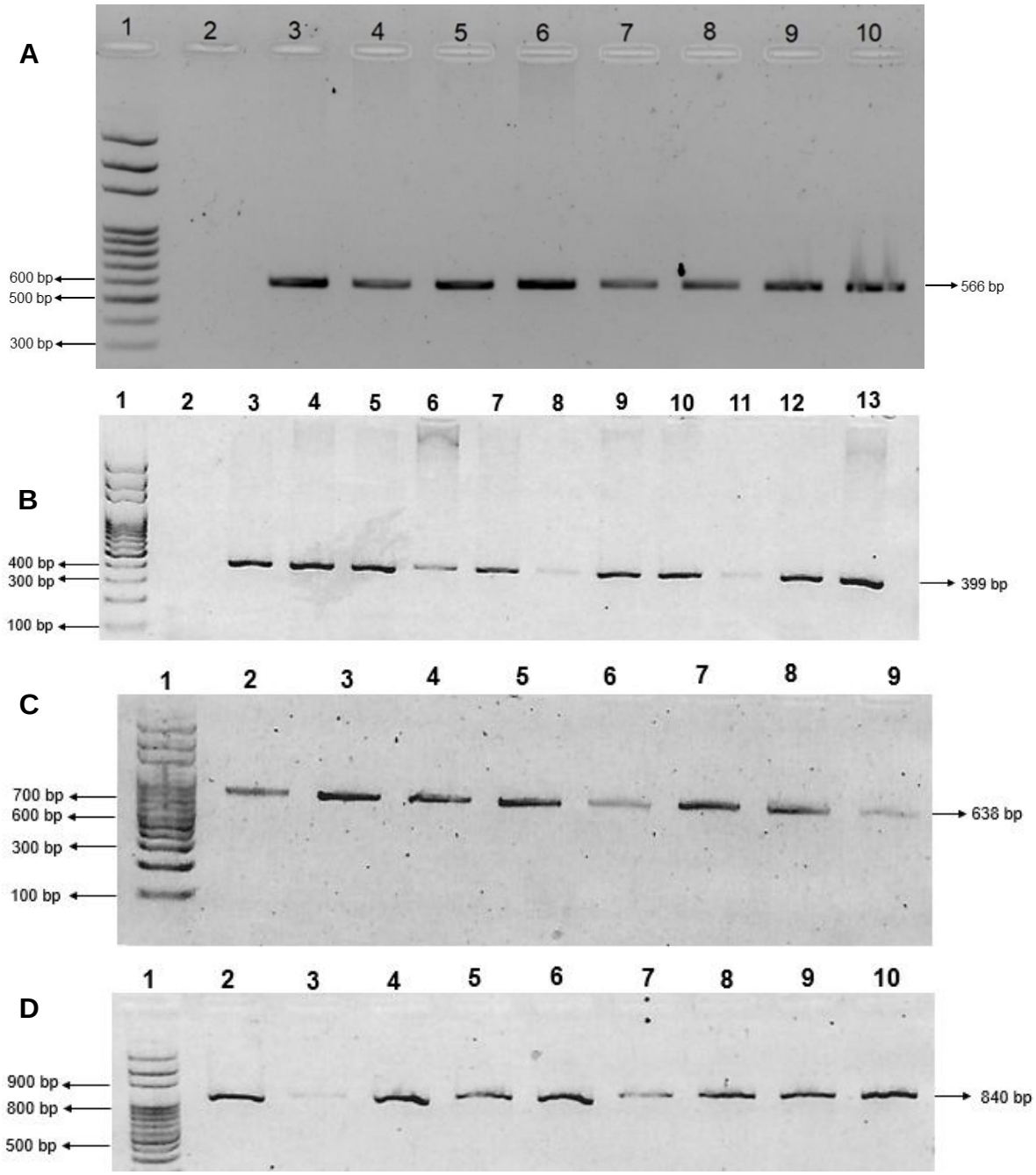
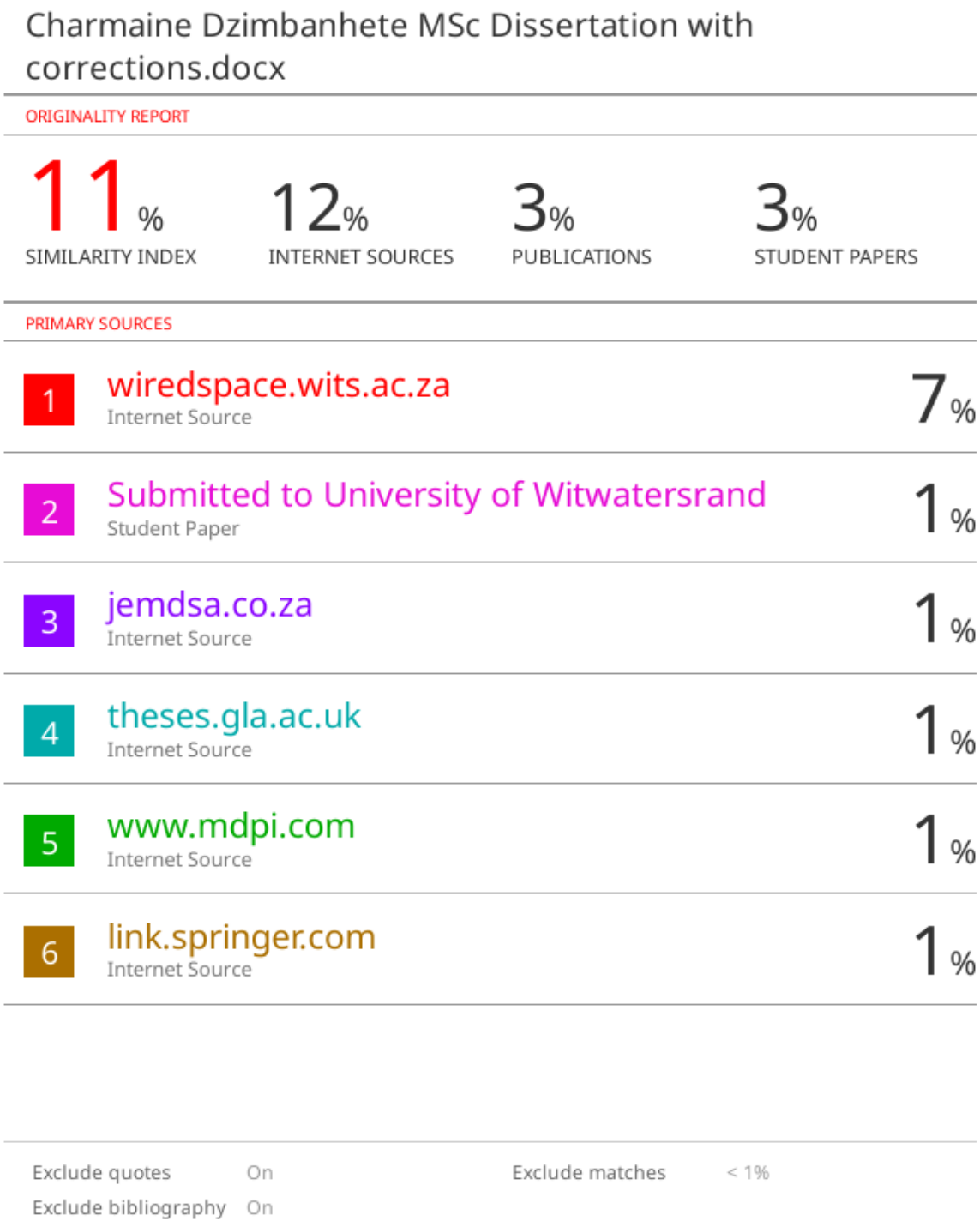


Figure 7.2: PCR products from amplification of all HLA-G exons run on a 1% agarose gel

A: Exon 1-2 PCR products: Lane 1: 100 bp DNA ladder; Lane 2: NTC; Lane 3-10: 566 bp PCR products of the *HLA-G* gene region flanking exon 1 and 2. **B:** Exon 3 PCR PCR products: Lane 1: 100 bp DNA ladder; Lane 2: NTC; Lane 3-13: 399 bp PCR products of the *HLA-G* gene region flanking exon 3. **C:** Exon 4-5 purified PCR products: Lane 1: 100 bp DNA ladder; Lane 2-9: Purified 638 bp PCR products of the *HLA-G* gene flanking exon 4 and 5. **D:** Exon 6-8 purified PCR products: Lane 1: 100 bp DNA ladder; Lane 2-10: Purified 840 bp PCR products of the *HLA-G* gene flanking exon 6, 7 and 8.

7.7 Appendix G: Turn-it-in report



Charmaine Dzimbanhete MSc Dissertation with corrections.docx

ORIGINALITY REPORT

11%	12%	3%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	wiredspace.wits.ac.za Internet Source	7%
2	Submitted to University of Witwatersrand Student Paper	1%
3	jemdsa.co.za Internet Source	1%
4	theses.gla.ac.uk Internet Source	1%
5	www.mdpi.com Internet Source	1%
6	link.springer.com Internet Source	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		