Association of Type 2 Diabetes Related SNPs with Predisposition of Gestational Diabetes Mellitus in Bangladeshi Women

Ph.D. Thesis

U.S. Mahzabin Amin

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY
FACULTY OF BIOLOGICAL SCIENCES
UNIVERSITY OF DHAKA
DHAKA, BANGLADESH

May, 2022

Association of Type 2 Diabetes Related SNPs with Predisposition of Gestational Diabetes Mellitus in Bangladeshi Women

By

U.S. Mahzabin Amin

A DISSERTATION SUBMITTED TO THE UNIVERSITY OF DHAKA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN BIOCHEMISTRY AND MOLECULAR BIOLOGY

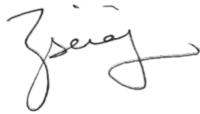
Faculty of Biological Sciences,
Department of Biochemistry and
Molecular Biology,
University of Dhaka,
Dhaka, Bangladesh.

Registration Number:
21/2015-16

Certificate

This is to certify that U.S. Mahzabin Amin has conducted her thesis work entitled, 'Association of Type 2 Diabetes Related SNPs with Predisposition of Gestational Diabetes Mellitus in Bangladeshi Women' under my supervision for the fulfilment of the degree of 'Doctor of Philosophy in Biochemistry and Molecular Biology' from the University of Dhaka. The work or any part of the thesis has not been submitted anywhere for any other degree.

Supervisor



……………………………

Date: 24/04/2022

Zeba Islam Seraj, Ph.D.
Professor, Department of Biochemistry and Molecular Biology
University of Dhaka.
Ph: +880-2-9661920x7647
zebai@du.ac.bd

Certificate

This is to certify that U.S. Mahzabin Amin has conducted her thesis work entitled, 'Association of Type 2 Diabetes Related SNPs with Predisposition of Gestational Diabetes Mellitus in Bangladeshi Women' under my supervision for the fulfilment of the degree of 'Doctor of Philosophy in Biochemistry and Molecular Biology' from the University of Dhaka. The work or any part of the thesis has not been submitted anywhere for any other degree.

Joint Supervisor

....

Date: 24/04/2022

Dr. Md. Salimullah
Director General (Additional charge),
National Institute of Biotechnology (NIB)
Ganakbari, Ashulia, Savar, Dhaka-1349
Tel: +88-02-7789458
salim2969@gmail.com

Dhaka University Institutional Repository

Dedication

To my dearest children
Mustaeen Ur Rahman and
Unaisa Sunayra Yashfee

Contents	Page No.
Abbreviation	xix
List of definitions	xxiii
Abstract	xxiv
1. Introduction	1
1.1 Background	2
1.1.1 Global scenario of Diabetes	2
1.1.2 Diabetes Mellitus: Bangladesh perspective	3
1.1.3 Women and Diabetes	4
1.1.4 Diabetes and pregnancy	5
1.2 Aim of this study	5
1.3 Objectives	5
1.4 Hypothesis	6
1.5 Structure of the Dissertation	6
Reference	7-8
2. Review of literature	9
2.1 Gestational Diabetes Mellitus (GDM)	10
2.1.1 Historical Background of GDM	10
2.1.2 Prevalence of GDM	10
2.1.3 Adverse outcome of GDM	11
2.1.4 Pathophysiology of GDM:	12
2.1.5 Genetic basis of GDM	14
2.1.6 Risk factors of GDM	14
2.2 Type 2 diabetes mellitus (T2DM):	14
2.2.1 Genetic pathophysiology of T2D	15
2.3 Diabetes in pregnancy (DIP)	15
2.4 GDM and the risk of diabetes	16
2.5 GDM and T2DM: share common pathophysiological Background	16
2.6 Single-nucleotide polymorphisms (SNPs)	17
2.7 CDK5 Regulatory Subunit Associated Protein 1-Like1(*CDKAL1)*	17
2.8 Fat mass and obesity associated gene (*FTO)*	18

Contents	Page No.
2.9 Heat Shock Protein Family A (Hsp70) Member 1 Like *(HSPA1L)*	20
2.10 Peroxisome proliferator-activated receptor gamma (*PPARG*)	21
2.11 Transcription factor 7 like 2 (*TCF7L2*)	22
2.12 Wolfram syndrome 1 (*WFS1*)	23
2.13 Case-control Studies	24
2.14 Odds ratio	25
2.15 Candidate polymorphism studies	26
2.16 HOMA indices in GDM	26
2.17 Hardy Weinberg Equilibrium (HWE)	27
2.18 Confounder	28
References	28-46
Methods and Materials	47
3.1 Study Design	48
3.1.1 Selection of target genes	48
3.1.2 Selection of target SNPs	49
3.1.3 Sample collection	50
3.1.4 Genotyping	52
3.1.5 Association analyses	52
3.2 Screening and enrollment of the participants	52
3.2.1 Recruitment of Study subjects	52
3.2.2 Eligibility Assessment	53
3.2.2.1 Inclusion criteria	53
3.2.2.2 Exclusion criteria	53
3.3 Diagnosis of GDM by Oral Glucose Tolerance Test (OGTT)	53
3.4 Insulin Indices	54
3.4.1 Fasting serum insulin	54
3.4.2 Homeostasis Model Assessment (HOMA)	54
3.4.3 Principles of the Procedure	55
3.5 Data collection Procedure	55
3.5.1 Demographic and anthropometric data collection	55
3.5.2 Body mass index	56

Contents	Page No.
3.5.3 Gravida	56
3.5.4 Occupational status	56
3.6 Blood Sample Collection	56
3.7 DNA Extraction	56
3.8 Quality assessment and quantification of DNA	57
3.9 Polymorphism analysis	58
3.9.1 Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP)	58
3.9.1.1 PCR amplification of regions spanning target SNPs	58
3.9.1.1.1 Primer Design	58
3.9.1.1.2 PCR amplification	59
3.9.1.1.3 Analysis of PCR Products	59
3.9.1.2 Restriction digestion	60
3.9.1.3 Analyses of Polymorphisms from gel electrophoresis of digestion products	60
3.9.1.3.1 Polyacrylamide gel electrophoresis (PAGE)	61
3.9.1.3.1.1 Preparation and procedure of 12% non-denaturing polyacrylamide gel electrophoresis (PAGE)	61
3.9.2 TaqMan allelic discrimination assay	62
3.9.3 Tetra-primer Amplification Refractory Mutation System Polymerase Chain Reaction (T-ARMS-PCR)	63
3.10 DNA sequencing	65
3.10.1 Template preparation	65
3.10.2 Cycle sequencing	65
3.10.3 Cycle sequencing PCR program	66
3.10.4 Purification of the post cycle sequencing product	66
3.10.5 Detection of the nucleotide by ABI PRISM® 3130 Genetic Analyzer	67
3.10.6 Sequence Data analysis	67
3.11 Allele and genotype frequency calculation	67
3.12 Statistical Analyses	68
3.13 Ethical consideration	69
References	69-71
4. Frequency Detection of the target SNPs related to T2D in women with and without GDM	72

Contents	**Page No.**
4.1 Study subjects	73
4.1.1 General characteristics of the study subjects	73
4.2 Frequency Detection of the Type 2 Diabetes (T2D) Related target SNPs	75
4.2.1 Frequency detection of the *CDKAL1* gene variants rs7756992 and rs7754840	75
4.2.1.1 General characteristics of the subjects genotyped for the polymorphisms	75
4.2.1.2 Genotype and allele frequencies of the *CDKAL1* gene variants	76
4.2.2 Frequency detection of the *FTO* gene variant rs8050136	78
4.2.3 Frequency detection of the *HSPA1L* gene variant rs2227956	79
4.2.4 Frequency detection of the *PPARG* gene variants rs3856806 and rs1801282	81
4.2.5 Frequency detection of the *TCF7L2* gene variants, rs12255372, rs10885406 and rs7903146	83
4.2.6 Frequency detection of the *WFS1* gene variant, rs10010131	86
4.3 Expected and observed genotype and allele frequencies of selected SNPs in the study participants	88
4.4 The frequency detection of the target SNPs in the study participants	89
4.5 Summary	90
References	91-92
5. Association of T2DM related SNPs with the predisposition of GDM	**93**
5.1 Association of T2DM linked genetic variants with GDM	94
5.1.1 Association of the *CDKAL1* gene variants rs7756992 and rs7754840 with GDM	95
5.1.2 Association of the *FTO* gene variant rs8050136 with GDM	96
5.1.3 Association of *HSPA1L* gene variant rs2227956 with GDM	97
5.1.4 Association of *PPARG* gene rs3856806 and rs1801282 variants with GDM	97
5.1.5 Association of the *TCF7L2* gene variant rs10885406 with GDM	99
5.1.6 Association of *TCF7L2* gene rs7903146 variant with GDM	100
5.1.7 Association of *TCF7L2* gene variant rs12255372 with GDM	101
5.1.8 Association of the *WFS1* gene variant rs10010131 with GDM	101
5.2 Association of the alleles of the selected T2DM linked genetic variants with GDM	102
5.3 Association haplotypes of T2DM linked genetic variants with GDM	104

Contents	**Page No.**
5.3.1 The Association of Haplotype of rs7756992 and rs7754840 with GDM	104
5.3.2 Association analyses of Haplotypes of *CDKAL1* and *HSPA1L* gene variants with GDM	104
5.3.3 The Association of the Haplotypes of the *PPARG* gene variants rs3856806 and rs1801282 with GDM	105
5.3.4 The Association of Haplotype of *TCF7L2* gene variants with GDM	106
5.4 Summary	107
Reference	108
6. Cumulative association of family history of diabetes and selected SNPs with the predisposition of GDM	109
6.1 Association of target SNPs and family history of Diabetes with GDM	110
6.1.1 Association of the *CDKAL1* Gene Variants with the Family History of Diabetes	110
6.1.2 Interaction between the *FTO* gene variant rs8050136 and family history of diabetes on the susceptibility of GDM	111
6.1.3 Interaction between the *HSPAL1* gene rs2227956 and family history of diabetes on the susceptibility of GDM	112
6.1.4 Interaction between the *PPARG* gene variants (rs3856806 and rs1801282) and family history of diabetes on the susceptibility of GDM	113
6.1.5 Association of the *TCF7L2* gene variant rs10885406 with the Family History of Diabetes	114
6.1.6 Association of the *TCF7L2* gene rs7903146 variant with the Family History of Diabetes	114
6.1.7 Association of the *TCF7L2* gene rs12255372 variant with the Family History of Diabetes	115
6.1.8 Association of the *WFS1* gene variant rs10010131 with the Family History of Diabetes	116
6.2 Summary	117
Reference	117-118
7. Cumulative association of target SNPs and Gravidity with GDM	119
7.1 Cumulative association of target SNPs and Gravidity with GDM	120
7.1.1 Association of the *CDKAL1* Gene Variants with Gravidity	120
7.1.2 Multigravidity and the *FTO* gene variant rs8050136 cumulatively increase chances of GDM	121
7.1.3 Multigravidity and the *HSPA1L* gene variant rs2227956 cumulatively increase chances of GDM	122

Contents	**Page No.**
7.1.4 Association of the *PPARG* gene variant rs3856806 and rs1801282 with Gravidity	123
7.1.5 Association of the *TCF7L2* gene variant rs10885406 with Gravidity	124
7.1.6 Association of the *TCF7L2* gene rs7903146 variant with Gravidity	124
7.1.7 Association of the *TCF7L2* gene variant rs12255372 with Gravidity	125
7.1.8 Association of the *WFS1* gene variant rs10010131 with Gravidity	126
7.2 Summary	126
Reference	127
8. Association of selected SNPs with anthropometric and metabolic parameters	128
8.1 Prevalence of GDM in the study participants	129
8.2 Fasting insulin level and insulin indices	130
8.3 Trimester	132
8.4 Age strata	133
8.5 BMI Groups	134
8.6 Summary	135
9. Discussion	136
9.1 Prevalence of GDM	137
9.2 Screening and Diagnosis of GDM	138
9.3 Genotype and allele distribution of the target SNPs	138
9.4 Selection of Confounder variables	139
9.5 Association of T2DM related SNPs with GDM	140
9.5.1 Association of *CDKAL1* gene variants	140
9.5.2 Association of *FTO* gene variant	141
9.5.3 Association of *HSPA1L* gene variant	141
9.5.4 Association of *PPARG* gene variants	141
9.5.5 Association of *TCF7L2* gene variants	142
9.5.6 Association of *WFS1* gene variant	143
9.6 Association analysis of haplotypes with GDM	143
9.7 Cumulative association of the target SNPs with GDM and family history of diabetes	144
9.8 Synergetic association of the target SNPs with GDM and gravidity	145

Contents	**Page No.**
9.9 General characteristics of the study population	145
9.9.1 Age	145
9.9.2 BMI	146
9.9.3 Gestation weeks	146
9.9.4 Fasting Insulin level and insulin indices	147
9.10 Power of the study	147
9.11 The strength and limitations of this study	148
Reference	148-157
10. Conclusion	158
10.1 Summary	159
10.2 Importance in terms of public health	160
10.3 Future Plan	161
Appendix	163

List of Tables

No. of table	Title of table	Page no
Table 2.1	List of maternal and fetal outcome of GDM	12
Table 3.1	List of genes used for SNP selection in this study	49
Table 3.2	List of selected SNPs and reported T2DM risk by their reference and altered alleles	50
Table 3.3	WHO 2013 Criteria	54
Table 3.4	List of primers used to amplify regions spanning variants genotyped by PCR-RFLP method	59
Table 3.5	List of primers used for sequencing	59
Table 3.6	PCR product size and restriction enzymes used for PCR-RFLP method	60
Table 3.7	List of primers used in T-ARMS method for the *FTO* gene rs8050136 genotyping	64
Table 4.1	General characteristics of the study participants	74
Table 4.2	Anthropometric and demographic data of the participants(n=468)	75
Table 4.3	Genotype and allele frequency of the *CDKAL1* gene variants (rs7756992 and rs7754840) in the study participants	77
Table 4.4	Genotype and allele frequency of *FTO* gene variant rs8050136 in the study participants	79
Table 4.5	Genotype and allele frequency of *HSPA1L* gene variant rs2227956 in the study participants	80
Table 4.6	Genotype and allele frequency of *PPARG* gene variants (rs3856806 and rs1801282) in the study participants	82
Table 4.7	Genotype and allele frequency of *TCF7L2* gene variant rs10885406 in the study participants	85
Table 4.8	Genotype and allele frequency of the *TCF7L2* gene variant rs7903146 in the study subjects	85
Table 4.9	Genotype and allele frequency of *TCF7L2* gene variants rs12255372 in the study subjects	86
Table 4.10	Genotype and allele frequency of *WFS1* gene variants rs10010131 in the study participants	87

No. of table	Title of table	Page no
Table 4.11	Genotype and allele frequency of the target SNPs in the study participants(n=505)	89
Table 5.1	Assessment of the risk of GDM provided by the studied variables	94
Table 5.2	Association of the allele frequencies of the target SNPs with GDM	103
Table 5.3	The association of haplotype of rs7756992 and rs7754840 with GDM	104
Table 5.4	The association of Haplotypes of *CDKAL1* and *HSPAL1* gene variants with GDM(n=466)	105
Table 5.5	The association of haplotype of rs3856806 and rs1801282 with GDM	106
Table 5.6	The association of haplotype of *TCF7L2* gene variants with GDM	106
Table 5.7	Nature of association of the target SNPs with GDM.	107

List of Figures

No. of Figures	Title of Figures	Page no
Figure 1.1	Distribution of diabetic people worldwide and per IDF Region in 2021–2045 adapted from International Diabetes Federation (IDF)(3)	3
Figure 1.2	Age-adjusted comparative prevalence (%) of diabetes (20–79 years) in the IDF South-East Asia Region in 2021	4
Figure 1.3	Number of deaths due to diabetes in adults by age and sex in 2021 adapted from IDF	4
Figure 2.1	Prevalence of GDM in Bangladesh	11
Figure 2.2	Schematic example of Case control Study	25
Figure 3.1	Steps of the overall study design	48
Figure 3.2	(a) GDM and T2DM have common pathophysiology adapted from (1) (b) Schematic diagram of selected genes variants of which related to the pathophysiology of T2DM	49
Figure 3.3	Flow chart of the overall methods followed in this study.	51
Figure 3.4	Screening for selecting study participants.	52
Figure 3.5	Measuring the concentration of DNA using a NanoDrop spectrophotometer	57
Figure 3.6	SNP genotyping by PCR-Restriction Fragment Length Polymorphism (PCR-RFLP)	58
Figure 3.7	Graphical representation of the TaqMan® Genotyping Assay (F.M. De La Vega et al. / Mutation Research 573, 2005)	62
Figure 3.8	Tetra-primer Amplification Refractory Mutation System-Polymerase Chain Reaction (T-ARMS-PCR) for SNP genotyping	64
Figure 3.9	Steps of DNA sequencing processes.	67
Figure 4.1	(a) RFLP analysis for genotyping. Lane 1(M): 1kb plus DNA ladder, lanes 2,4,6,8,10,12,14 and 16; Undigested PCR product (C=control); lane 3, 5, 7 and 15; AA Homozygous genotype, lanes 9, 11 and 13; AG Heterozygous genotypes and lane 17; GG Homozygous genotype. (b) Confirmation of the RFLP results by DNA sequencing.	76
Figure 4.2	A single run of 30 samples yielded an allele discrimination plot that comprised both cases and controls, as well as representative amplification plots for each genotype and a negative control. Negative controls were represented by black squares, while GG, GC, and CC genotypes were represented by blue, green, and red dots, respectively.	77

No. of Figures	Title of Figures	Page no
Figure 4.3	(a) Lane1 (M): 1Kb+ DNA Ladder. Lane 1, 2, 4, 5, 6, 9, 10, 12, 14, 16 homozygous for wild type (CC). Lane 3, 8, 13, and 15 are heterozygous (AC). Lane 7 and 11 are homozygous for mutant allele (AA). (b)Validation of the T-ARMS results by DNA sequencing.	78
Figure 4.4	(a) Lane 1(M): 1kb plus DNA ladder, lanes 2,4,5,7,9,11,12,15 and 16: TT homozygous; lanes 6,8 and 14: CT heterozygous; lane 13 CC homozygous genotypes. (b) Confirmation of rs2227956 genotypes by DNA sequencing	80
Figure 4.5	(a) Lane 1 (M): 1kb plus DNA ladder, lanes 2,4,6,8,10,12,14, and 16; (C=Control) PCR product(undigested); lane 3,7,9,13,15 and 17; CC Homozygous, lanes 5 and 11; CT genotypes. (b) Confirmation of rs3856806 genotypes by DNA sequencing	81
Figure 4.6	An allele discrimination plot generated from a single run of 30 samples, which included both cases and controls, as well as representative amplification plots for each genotype and a negative control. GG, GC, and CC genotypes were represented by blue, green, and red dots, respectively; negative controls were represented by black squares.	82
Figure 4.7	Lane 1 and 40 (M): 1kb plus DNA ladder; lanes 2-39 digested PCR products for rs12255372 genotyping	83
Figure 4.8	Confirmation of the rs12255372 of *TCF7L2* gene by sequencing randomly selected samples	84
Figure 4.9	Allele discrimination plots (a) rs10885406 and (b) rs7903146 from two single runs of 35 samples containing both controls and cases and amplification plots for each genotype, as well as a negative control. Blue, green and red dots denoted in (a) GG, AG and AA genotypes; and in (b) TT, CT and CC genotypes respectively; black squares showed negative controls in both plots.	84
Figure 4.10	(a) Lane 1(M): 1kb plus DNA ladder, lanes 2 and 12: AA homozygous; lanes 3,4,6,7,8,10,11,13-18,20 and 24: GG homozygous; lanes 5,9,19 and 21-23: AG heterozygous genotypes. (b) Confirmation of the RFLP results by DNA sequencing.	86
Figure 4.11	Comparison of expected and observed genotype frequencies of the target SNPs in control and cases.	88
Figure 4.12	Comparison of expected and observed allele frequencies of the target SNPs in control and cases.	89
Figure 5.1	Associations of rs7756992(AA/AG/GG) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed circles and whiskers representing the 95% confidence intervals. (*, **, *** $P<0.05$, $P<0.01$, $P<0.001$)	95

No. of Figures	Title of Figures	Page no
Figure 5.2	Associations of rs7754840(GG/GC/CC) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed squares and whiskers representing the 95% confidence intervals. (*, **, *** $P<0.05$, $P<0.01$, $P<0.001$)	96
Figure 5.3	Associations of rs8050136 (CC/AC/AA) with GDM under different genetic models adjusted for family history of diabetes and gravidity.	97
Figure 5.4	Associations of *HSPA1L* gene variant rs2227956 (TT/CT/CC) with GDM under different genetic models adjusted for family history of diabetes and gravidity	97
Figure 5.5	Associations of the *PPARG* gene rs3856806 (CC/CT/TT) with GDM under different genetic models adjusted for family history of diabetes and gravidity	98
Figure 5.6	Associations of *PPARG* gene rs1801282 (CC/CG/GG) with GDM under different genetic models adjusted for family history of diabetes and gravidity	99
Figure 5.7	Associations of the *TCF7L2* gene rs10885406 (AA/AG/GG) with GDM under different genetic models adjusted for family history of diabetes and gravidity	100
Figure 5.8	Associations of *TCF7L2* gene rs7903146 (CC/CT/TT) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed squares and whiskers representing the 95% confidence intervals.	100
Figure 5.9	Associations of *TCF7L2* gene rs12255372 (GG/GT/TT) with GDM under different genetic models adjusted for family history of diabetes and gravidity. (*, **, *** $P<0.05$, $P<0.01$, $P<0.001$)	101
Figure 5.10	Associations of rs10010131 (GG/AG/AA) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed squares and whiskers representing the 95% confidence intervals.	101
Figure 6.1	Correlation between GDM Risk with cumulation of the *CDKAL1* gene variant rs7756992 polymorphism, and family history of diabetes. (*, **, *** $P<0.05$, $P<0.01$, $P<0.001$)	110
Figure 6.2	Correlation between GDM Risk with cumulation of the *CDKAL1* rs7754840 polymorphism, and family history of diabetes under four genetic models adjusted for gravidity.	111
Figure 6.3	Correlation between GDM Risk with cumulation of the *FTO* gene rs8050136 polymorphism, and family history of diabetes under different genetic models and adjusted for gravidity.	112
Figure 6.4	Association between GDM Risk with cumulation of the *HSPAL1* gene rs2227956 polymorphism, and family history of diabetes.	112

No. of Figures	Title of Figures	Page no
Figure 6.5	Cumulative association of GDM Risk of the *PPARG* gene polymorphisms, and family history of diabetes.	113
Figure 6.6	Cumulative Correlation between GDM Risk of the *TCF7L2* gene rs10885406 polymorphism, and family history of diabetes. (*, **, *** *P*<0.05, *P*<0.01, *P*<0.001)	114
Figure 6.7	Cumulative correlation between GDM Risk of the *TCF7L2* gene rs7903146 polymorphism, and family history of diabetes.	115
Figure 6.8	Cumulative association of GDM Risk with the *TCF7L2* gene rs12255372 polymorphism, and family history of diabetes.	116
Figure 6.9	Cumulative association of GDM Risk with the *WFS1* gene rs10010131 polymorphism, and family history of diabetes.	116
Figure 7.1	Association of GDM Risk with cumulation of the *CDKAL1* gene variants (a) rs7756992 (b) rs7754840, and gravidity.	120
Figure 7.2	Correlation between GDM Risk with cumulation of the *FTO* gene rs8050136 polymorphism, and gravidity. (*, **, *** P<0.05, P<0.01, P<0.001)	122
Figure 7.3	Correlation between GDM Risk with cumulation of the *HSPA1L* gene rs2227956 polymorphism, and gravidity.	122
Figure 7.4	Association of GDM Risk with cumulation of the *PPARG* gene polymorphisms, and gravidity.	123
Figure 7.5	Correlation between GDM Risk with cumulation of the *TCF7L2* gene rs10885406 polymorphisms, and gravidity.	124
Figure 7.6	Association of GDM Risk with cumulation of the *TCF7L2* gene rs7903146 polymorphisms, and gravidity.	125
Figure 7.7	Interaction between GDM Risk with cumulation of the *TCF7L2* gene rs12255372 polymorphism, and gravidity.	125
Figure 7.8	Cumulative association of GDM Risk with the *WFS1* gene variant rs10010131, and gravidity.	126
Figure 8.1	Fasting insulin levels between control and GDM groups in each genotype of the target SNPs	130
Figure 8.2	Fasting insulin levels between control and GDM groups in each genotype of the target SNPs;1: *CDKAL1* rs7754840; 2: *CDKAL1* rs7756992; 3: *FTO* rs8050136; 4: *HSPA1L* rs2227956; 5: *PPARG* rs3856806; 6: *PPARG* rs1801282; 7: *TCF7L2* rs10885406; 8: *TCF7L2* rs7903146; 9: *TCF7L2* rs12255372; 10: *WFS1* rs10010131.	131
Figure 8.3	Percentage of total diagnosed GDM in different trimesters	132
Figure 8.4	Comparison of study variables between cases and controls within three trimesters.	133
Figure 8.5	Association of SNPs with GDM in different age groups.	134
Figure 8.6	The changes in odds of GDM with the increase in BMI of the participants	134
Figure 8.7	Association of the target SNPs with GDM within the BMI group ≥30 kg/m^2. a: rs3856806, b: rs1801282, c: rs8050136, d: rs7756992, e: rs7754840, f: rs2227956, g: rs10010131; h: rs12255372, i: rs7903146, j: rs10885406. (*, **, *** *P*<0.05, *P*<0.01, *P*<0.001)	135

Abbreviation

A	
A	Adenine
ATP	Adenosine tri phosphate
B	
bp	Base pair
BMI	Body Mass Index
BSMMU	Bangabandhu Sheikh Mujib Medical University
C	
C	Cytosine
⁰C	Degrees Celsius
CDKAL1	CDK5 Regulatory Subunit Associated Protein 1-Like1
CGAS	Candidate Gene Association Studies
cm	Centimeter
D	
ddH2O	Distilled deionized water
DIP	Diabetes in pregnancy
DM	Diabetes Mellitus
DNA	Deoxyribonucleic acid
E	
e.g.	For example
et al.	And others
EDTA	Ethylene diamine tetra acetic acid
EtBr	Ethidium bromide
etc	Etcetera
F	
FHD	Family history of diabetes
FPG	Fasting Plasma Glucose
FTO	Fat mass and obesity associated gene
G	
G	Guanine
GDM	Gestational Diabetes Mellitus
GIPR	Glucose-dependent insulinotropic polypeptide

GLP1R	Glucagon-like peptide-1
GWAS	Genome wide association study
H	
HIP	Hyperglycemia in pregnancy
hr	Hour
HWE	Hardy-Weinberg equilibrium
HSPA1L	Heat Shock Protein Family A (Hsp70) Member 1 Like
I	
IADPSG	The International Association of the Diabetes and Pregnancy Study Groups
IDF	International Diabetes Federation
IGF-1	insulin-like growth factor 1
IR	Insulin resistance
IRX3	Iroquois homeobox 3
IRX5	Iroquois homeobox 5
J	
K	
kb	kilo base
L	
LD	Linkage disequilibrium
M	
M	Molar
MAF	Minor allele frequency
mg	Milligram
min	Minute
ml	Milliliter
mM	Millimolar
N	
NCD	Non-communicable disease
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ng	Nanogram
No./no.	Number

O	
OGTT	Oral Glucose Tolerance test
OPG	Plasma glucose level one hour after glucose load
OR	Odds Ratio
P	
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PPARG	Peroxisome proliferator-activated receptor gamma
R	
RFLP	Restriction Fragment Length Polymorphism
rpm	Rotation per minute
RT	Room temperature
S	
SEA	South East Asia
SD	Standard Deviation
T	
T	Thymine
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
T-ARMS-PCR	Tetra-primer Amplification Refractory Mutation System Polymerase Chain Reaction
TE	Tris EDTA
TPG	Plasma glucose level one hour after glucose load
TCF7L2	Transcription factor 7 like 2
Tm	Melting temperature
U	
U	Enzyme unit
µg	Microgram
µl	Microliter
UV	Ultraviolet
V	
vol	Volume
vs.	Versus
V	Volt

W	
WFS1	Wolfram syndrome 1
WFS	Wolfram syndrome
WHO	World Health organization

List of Definitions

Allele frequency- The number of times the allele of interest is observed in a population is divided by the total number of copies of all the alleles at that particular genetic locus in the population is defined as an allele frequency.

Hardy Weinberg Equilibrium (HWE)-The Hardy-Weinberg (HW) principle refers to the connection between allele frequencies and counts of genotype in successive generations in the absence of any disturbing factors.

Predisposition-A genetic predisposition (sometimes also called genetic susceptibility) is an increased likelihood of developing a particular disease based on a person's genetic makeup.

Gravidity- Gravidity is defined as the number of times that a woman has been pregnant. For example, a woman who is described as 'gravida 2 (sometimes abbreviated to G2) has had two pregnancies.

Primigravida- an individual pregnant for the first time.
Multigravida-A multigravida has been pregnant more than once.

Abstract

Background: Recent advances in genetic studies have revealed a number of susceptible loci for Type 2 diabetes mellitus (T2DM). In this study, we attempted to analyze the independent effect of variants in some of these loci on Gestational Diabetes Mellitus (GDM). The association of single nucleotide polymorphisms (SNPs) with the susceptibility of GDM was studied in a group of Bangladeshi women.

Methods: Ten T2DM-related SNPs from six loci were selected. In this case-control study, 219 subjects with GDM and 286 subjects with normoglycemic controls were genotyped for the selected SNPs by PCR-RFLP, T-ARMS, and TaqMan™ allelic discrimination assay methods. Genotyping results were confirmed by DNA sequencing and replicated TaqMan™ assay. We analyzed the allele and genotype distribution between the cases and controls. The associations between SNPs and GDM were examined by logistic regression with five different genetic models adjusted for family history of diabetes (FHD) and gravidity. The cumulative associations of the target SNPs and the confounding variables with GDM were analyzed by interaction analyses.

Results: We examined the effects of SNPs from CDK5 Regulatory Subunit Associated Protein 1-Like1(*CDKAL1)*, Fat mass and obesity associated gene (*FTO)*, Heat Shock Protein Family A (Hsp70) Member 1 Like (*HSPA1L),* Peroxisome proliferator-activated receptor gamma (*PPARG)*, Transcription factor 7 like 2 (*TCF7L2)*, and Wolfram syndrome 1 (*WFS1*) on the risk of GDM, with odds ratios ranging from 0.58 to 2.09. The *CDKAL1* variants, rs7756992 (OR=1.6, *P*=0.02) and rs7754840 (OR=2.09, *P*=0.047), and the *TCF7L2* rs12255372 (OR= 1.44, *P*=0.046) were significantly associated with the susceptibility of GDM. However, no significant association was detected between SNPs from *FTO*, *HSPA1L*, *PPARG*, and *WFS1* with GDM. The risk alleles containing (CG) haplotype of the *CDKAL1* gene variants, rs7756992 and rs7754840, conferred significant (*P*=0.032) disease susceptibility with an odds ratio of 1.43 (1.03-1.98). Concomitant presence of the risk alleles of these SNPs and positive FHD in any pregnant woman increased the chance of developing GDM by 1.5 to 4.8 folds. Significant increase in the susceptibility of GDM resulted from the *CDKAL1* rs7756992 (OR=3.08, *P*=0.038) and *TCF7L2* rs10885406 (OR=3.42, *P*=0.015). The synergistic effect of risk alleles of these SNPs and multigravidity increased the odds of

GDM by more than 1.5 folds in different genotypes, but a significant increase was revealed from the interaction analyses for *FTO* rs8050136 (*P*=0.0068).

Conclusion: Several SNPs related to T2DM were associated with the risk of GDM through genotypic effects alone or interactions with the family history of diabetes and gravidity. These findings do not indicate to a single significant T2DM gene linked to GDM, but they do support the idea that T2DM is causally linked to GDM through several T2DM susceptibility genes and interactions with other factors. These associations also provided the possibility of potential markers for prediction of GDM and T2DM in Bangladeshi women.

1.Introduction

1.1 Background

Gestational diabetes mellitus (GDM) is defined as a metabolic disorder observed during pregnancy which usually disappears after childbirth. Due to maternal hyperglycemia, GDM leads to adverse maternal and fetal outcomes (1). Women diagnosed with GDM have a higher risk of progression to type 2 diabetes mellitus (T2DM) (2). Determination of the high-risk population by association of single nucleotide polymorphisms (SNPs) related to T2DM and correlating with GDM can aid in effective strategies for preventing the onset of T2DM. Bangladesh is ranked as being the eighth-highest country in the world in the number of people afflicted with Diabetes in 2021(3). In last few years, increasing frequency of GDM has been observed in Bangladesh (4). Therefore, this study has focused on some T2DM related variants and their association with GDM that can suggest lifestyle intervention of pre-disposed individuals to prevent occurrence of GDM in subsequent pregnancies as well as eventually succumbing to T2DM.

1.1.1 Global scenario of Diabetes

The epidemic of diabetes which is one of the major non-communicable diseases, poses a significant threat to global public health. The prevalence of this disease has rapidly increased worldwide over the past few decades (5, 6). The number of diabetic individuals increased from 108 million in 1980 to approximately 537 million (1 in 10 persons) in 2021 and may reach 783 million by 2045 (3, 7). In low- and middle-income countries, the prevalence of this disease has been rising more rapidly than in high-income countries (8). Estimating the prevalence of different types of diabetes, i.e., Type 1 Diabetes (T1D) and Type 2 Diabetes (T2D) separately, is difficult due to the lack of independent investigations (9). Nevertheless, one estimate shows that approximately 87-91% and 7-12% of all diabetes cases in developed countries are T2D and T1D, respectively (9).

In addition to that, other types of diabetes are reported to affect 1 to 3% of the total diabetes cases (9). Prolonged diabetes can damage the heart, blood vessels, eyes, kidneys, nerves and cause a two- to three-fold elevated risk of heart attacks and strokes in adults (10-12). Many people are unaware of their physical conditions and remain undiagnosed worldwide, especially in economically disadvantaged regions. In 2021,

almost one in two diabetic adults (20–79 years old) were reported to be unaware of their diabetes status, and the total number of them is approximately 239.7 million (44.7%) (13). The proportion of undiagnosed diabetes is the third highest (51.2%) in South East Asia (SEA) (3). It also contributes significantly to the predicted decline in life expectancy. Diabetes stood ninth as the leading cause of death in 2019 which estimated 1.5 million deaths directly caused by this disease and over 80% of which occur in developing countries (8).

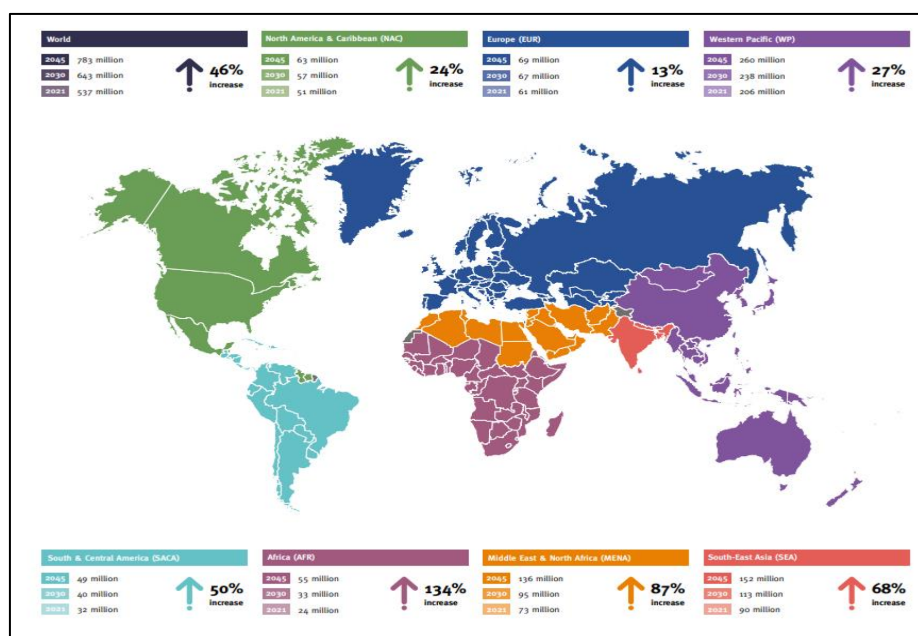


Figure 1.1: Distribution of diabetic people worldwide and per IDF Region in 2021–2045 adapted from International Diabetes Federation (IDF)(3)

1.1.2 Diabetes Mellitus: Bangladesh perspective

In Bangladesh, prevalence of non-communicable diseases (NCD) is increasing day by day compared to the communicable diseases lading to an epidemiological transition (14). Diabetes mellitus, notably T2D, is now recognized as a major chronic public health problem. The prevalence of this disease is the highest among SEA countries (Figure 1.2). In 2021 IDF estimated 13.1 million people in our country with this disease with the anticipated number of people increasing to 22.3 million in 2045 (3). Nearly half of the population with diabetes do not know that they have diabetes and therefore do not receive any treatment. It was noted that 43.5% (5.7 million) of cases of diabetes was undiagnosed in 2021 (3). During the 90s, the number of diabetes-affected

population was low. In 1995 it was only 4% increasing to 5% in 2000 and to 9% in 2006 - 2010 period. A 2.5 folds increase of diabetes was observed in the last two decades, which was 4.0% in 1995–2000 and 10.4% in 2010–2019 (15). Alarmingly, many children and young people develop diabetes, which is a grave concern for any country (16). Bangladesh is one of the countries with the lowest diabetes-related annual expenditure (USD 77) per person (3).

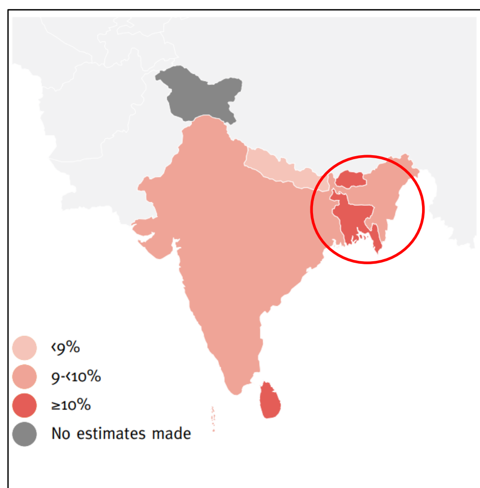


Figure 1.2: Age-adjusted comparative prevalence (%) of diabetes (20–79 years) in the IDF South-East Asia Region in 2021(3)

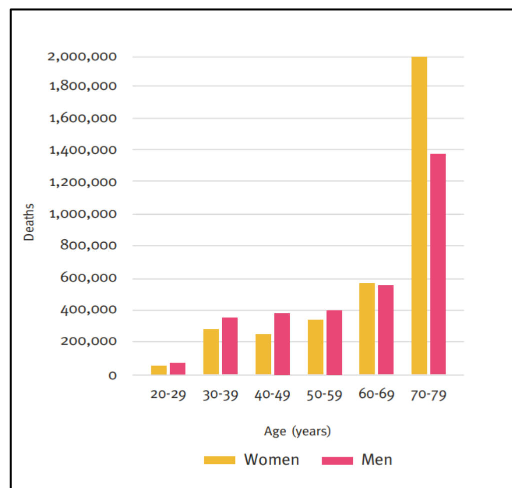


Figure 1.3: Number of deaths due to diabetes in adults by age and sex in 2021 adapted from IDF (3)

1.1.3 Women and Diabetes

An estimated 223 million women (20-79 years) live with diabetes, which may increase to 343 million by 2045 (3). The prevalence of this disease was 10.2% in women aged 20–79 years in 2021 which is slightly lower (10.8%) than men. The pooled prevalence was slightly lower among women than men (6.70% vs. 7.34%) (15). Even though diabetes affects both men and women equally, the repercussions of diabetes are more severe in women. In addition, compared to men, women have fewer opportunities to be treated, less access to care, and less support to deal with this situation. The risks of diabetes-related complications like heart disease, blindness, kidney disease, and depression are higher in women; specifically, the risk of the most common complication, heart disease, is increased by four times (17-19). Pre-menopausal diabetic women are 50% more likely to die from heart disease than men (20). In 2021, the number of deaths due to diabetes was higher in women of age groups 60-69 and 70-79 years compared to men (Figure 1.3) (3).

1.1.4 Diabetes and pregnancy

Hyperglycemia during pregnancy is detrimental to maternal and fetal health. In 2021, 16.7% (21.1 million) live births had some form of hyperglycemia in pregnancy (HIP). Among those, 80.3% were due to GDM, while 10.6% were the result of diabetes detected before pregnancy, and 9.1% were due to diabetes (including T1D and T2D) first noticed in pregnancy. The proportion of HIP was the highest in SEA among IDF Regions at 25.9%, and one in four live births were affected in 2021 (3). GDM is characterized by elevated blood sugar levels detected in pregnancy. Many maternal and fetal co-morbidities have been linked with GDM. GDM increases the risk of developing T2D for both mother and offspring later in life. In addition, GDM has also been linked with cardiovascular disease and metabolic syndrome (21).

1.2 Aim of this study

Due to the adverse effect of GDM on mothers and their children, studies of the complex etiology of GDM are need of time to reduce the occurrence of GDM. In addition, studies on the maternal health problems during the time of conception are also needed to reduce the rate of GDM. This metabolic complication of pregnancy leads to adverse health effects, both short and long terms, in mother and fetus. It shows an alarming prospect due to this increase in GDM and related health effects. Moreover, there is very limited number of research that explored the relation between these conditions. In addition, there is little or no knowledge about the genetic basis of this disease as well as of T2DM in our population. Recent genome-wide association studies (GWAS) reported a number of genetic variants related to T2DM in different populations (22). Thus, this study aims to develop a better understanding of the association of some T2DM related genetic variants with GDM in our population and explore whether and to what extent the presence of these variants increases the risk of GDM in pregnancies and T2DM later in life.

1.3 Objectives

The overall objective is to identify the association of T2DM related SNPs with predisposition of GDM. The specific objectives are:

1. To detect the frequency of selected SNPs related to T2D in women with and without GDM

2. To identify the association of these SNPs with GDM.
3. To check the association of tested polymorphisms in women with GDM and their positive family history of T2DM among 1st degree relatives.

1.4 Hypothesis: SNPs related with T2DM are associated with the predisposition of GDM

1.5 Structure of the Dissertation This dissertation consists of ten chapters and references.

Topic	Name of the chapter	Content
Introduction	Chapter 1	Introduces the problems related to diabetes and GDM. It also describes the aims and objectives of the study.
Review of Literature	Chapter 2	Provides the literature review regarding this study e.g., overview of prevalence, etiology, pathophysiology, and risk factors of GDM and T2DM
Methods	Chapter 3	Describes the methods and materials used in this study
Results	Chapter 4	Presents the findings of the first specific objective, which is to determine the frequency of selected SNPs in control and GDM groups.
	Chapter 5	Presents the findings of the second objective, which is to determine the association between T2D related SNPs and GDM.
	Chapter 6	Shows findings of the third objective, which is to determine the association between family history of diabetes and GDM.
	Chapter 7	Describes the findings of the association between gravidity and GDM.
	Chapter 8	Describes findings of the association analyses of selected SNPs with anthropometric and metabolic parameters
Discussion	Chapter 9	Discusses the overall findings
Conclusion	Chapter 10	Concludes the results obtained from this study, describes the public health-related significance and suggests future research.
References	References	At the end of each chapter

References

1. Jarmuzek P, Wielgos M, Bomba-Opon D. Placental pathologic changes in gestational diabetes mellitus. Neuro endocrinology letters. 2015;36(2):101-5.

2. Ward RJ, Fryer AA, Hanna FW, Spencer N, Mahmood M, Wu P, et al. Inadequate postpartum screening for type 2 diabetes in women with previous gestation diabetes mellitus: A retrospective audit of practice over 17 years. International journal of clinical practice. 2021;75(9):e14447.

3. International Diabetes Federation. IDF_Atlas_10th_Edition_2021.pdf.

4. Nusrat-Sultana HM, Sharmin-Jahan M-H, Sandesh-Panthi Y-A, Fariduddin MJStJ. Alarming frequency of Gestational Diabetes Mellitus (GDM) attending a tertiary care hospital in Bangladesh. 2016.

5. World Health Organization. Global report on diabetes. Part 1: Global burden of diabetes.; 2016.

6. Polonsky KS. The past 200 years in diabetes. The New England journal of medicine. 2012;367(14):1332-40.

7. Collaboration NCDRF. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet (London, England). 2016;387(10027):1513-30.

8. World Health Organization. Diabetes 2021 [updated 10 November 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/diabetes.

9. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes research and clinical practice. 2018;138:271-81.

10. Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, Di Angelantonio E, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet (London, England). 2010;375(9733):2215-22.

11. Study† GBaVICobotVLEGotGBoD. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. The Lancet Global health. 2021;9(2):e144-e60.

12. Saran R, Li Y, Robinson B, Ayanian J, Balkrishnan R, Bragg-Gresham J, et al. US Renal Data System 2014 Annual Data Report: Epidemiology of Kidney Disease in the United States. American journal of kidney diseases : the official journal of the National Kidney Foundation. 2015;66(1 Suppl 1):Svii, S1-305.

13. Ogurtsova K, Guariguata L, Barengo NC, Ruiz PL, Sacre JW, Karuranga S, et al. IDF diabetes Atlas: Global estimates of undiagnosed diabetes in adults for 2021. Diabetes research and clinical practice. 2022;183:109118.

14. Chowdhury R, Alam DS, Fakir, II, Adnan SD, Naheed A, Tasmin I, et al. The Bangladesh Risk of Acute Vascular Events (BRAVE) Study: objectives and design. European journal of epidemiology. 2015;30(7):577-87.

15. Akhtar S, Nasir JA, Sarwar A, Nasr N, Javed A, Majeed R, et al. Prevalence of diabetes and pre-diabetes in Bangladesh: a systematic review and meta-analysis. 2020;10(9):e036086.

16. Biswas T, Islam A, Rawal LB, Islam SM. Increasing prevalence of diabetes in Bangladesh: a scoping review. Public health. 2016;138:4-11.

17. Centers for Disease Control and Prevention. Diabetes and Women 2020 [Available from: https://www.cdc.gov/diabetes/library/features/diabetes-and-women.html.

18. Villar E, Chang SH, McDonald SP. Incidences, treatments, outcomes, and sex effect on survival in patients with end-stage renal disease by diabetes status in Australia and New Zealand (1991 2005). Diabetes care. 2007;30(12):3070-6.

19. Shea L, Owens-Gary MJDV. Diabetes and depression in older women–double the risk, double the burden. 2009;54:8-11.

20. Huxley R, Barzi F, Woodward M. Excess risk of fatal coronary heart disease associated with diabetes in men and women: meta-analysis of 37 prospective cohort studies. BMJ (Clinical research ed). 2006;332(7533):73-8.

21. Coustan DR. Gestational diabetes mellitus. Clinical chemistry. 2013;59(9):1310-21.

22. Cirillo E, Kutmon M, Gonzalez Hernandez M, Hooimeijer T, Adriaens ME, Eijssen LMT, et al. From SNPs to pathways: Biological interpretation of type 2 diabetes (T2DM) genome wide association study (GWAS) results. PloS one. 2018;13(4):e0193515-e.

2. Review of Literature

2.1 Gestational Diabetes Mellitus (GDM)

Gestational Diabetes Mellitus (GDM) is defined as the varying degree of carbohydrate intolerance that is diagnosed during pregnancy for the first time and is a common obstetric complication. It affects one in six births worldwide (1). This disorder is defined by an inability to compensate for pregnancy-induced insulin resistance by increasing insulin secretion. A number of studies have found that people with GDM have a higher likelihood of having T2DM in future, as well as short and long-term effects on their offspring's metabolic health. (2).

2.1.1 Historical Background of GDM

Women with diabetes mellitus (DM) experienced poor pregnancy outcomes during the past century. It was observed in the 1940s that women who had DM later in life had an unusually high rate of neonatal mortality and unfavorable pregnancy outcomes (3). GDM was first defined in the 1950s as a transient maternal disease that had a deleterious impact on fetal outcomes and then healed after delivery (4). It was discovered in the 1960s the intensity of glucose intolerance experienced during pregnancy was linked to the woman's chance of acquiring diabetes after about a few years of giving birth. The oral glucose tolerance test (OGTT) was given a new interpretation. (5). Modern methods for measuring blood glucose adapted the cut-off values of the OGTT during the 1980s (6). The definition of GDM was set during the 'Fourth International Workshop-Conference on Gestational Diabetes' in 1998 (7).

2.1.2 Prevalence of GDM

Prevalence of GDM reflects the background rate of T2DM in the respective population. Along with the epidemic of diabetes the prevalence of GDM has increased worldwide and occurs in 1 to 28% of all pregnancies varying substantially between population and diagnostic criteria used(8). In our country prevalence of this disease also increasing rapidly. In 2015 Sandesh et al reported 30% and 31.88% prevalence by using WHO 2013 and 1999 criteria respectively (9). This growing prevalence rate is alarming as there are a number of adverse outcomes of GDM results both in mother and fetus. It affects approximately 5-10 % of pregnant women in Asia and 1-3% in Western countries (10, 11). This complication is growing substantially in the prevalence of

36.6% of total pregnancies in Bangladesh (9, 12). This prevalence is estimated from 0.7 to 51% in Asia (8, 13-15). The huge disparity in prevalence rates could be explained by differences in ethnicity, diagnostic criteria, screening methodologies, and population characteristics. (8).

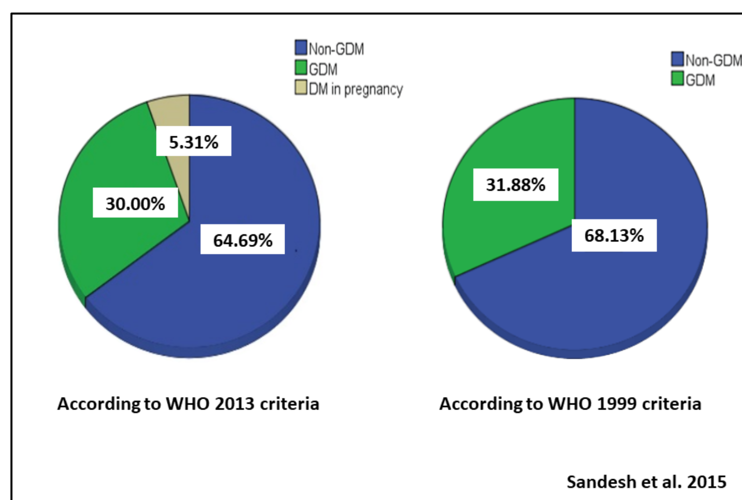


Figure 2.1: Prevalence of GDM in Bangladesh

2.1.3 Adverse outcome of GDM:

GDM can cause large-for-gestational-age newborns, increasing the risk of pregnancy and birth difficulties for both the mother and the baby. Maternal outcomes include preaclamsia, premature delivery, increased risk of developing diabetes after delivery etc on the other hand macrosomia, stillbirth, congenital malformation and long-term complications like obesity and diabetes are some fetal outcomes (**Table 2.1).**

Table 2.1: List of maternal and fetal outcome of GDM

Maternal Outcomes	Fetal outcomes
• Preeclampsia(16)	• Macrosomia (17, 18)
• Hypertension	• Shoulder dystocia or birth injury (17, 18)
• Premature delivery	• Perinatal mortality which includes stillbirth and early neonatal death (17, 18)
• Urinary and genital tract infections	• Congenital malformation
• Polyhydramnios	• Neonatal hypoglycemia(19)
• Increased risk of cesarean delivery(16)	• Polycythemia (17, 18)
• Increased risk of developing diabetes after pregnancy	• Long-term complications: ▪ Increased risk of glucose intolerance(20) ▪ Diabetes(21) ▪ Obesity(22)

2.1.4 Pathophysiology of GDM:

The carbohydrate metabolism changes progressively during pregnancy to satisfy the rising demands of the mother and the growing fetus. For production of endogenous glucose, liver is the main source in a non-pregnant woman. For fasting glucose, the average plasma concentration is ~ 5.0 mmol/l, which ensures a equilibrium between consumption and production (23). During progression of pregnancy fasting glucose level drops with the increase in hepatic glucose levels (24). Normally, the production of hepatic glucose is restrained by insulin, but in pregnancy, though there is an increase in fasting insulin concentration, the hepatic glucose production upsurges. (25). As the result of this situation, maternal hepatic insulin sensitivity decreases that leads to a reduced suppression of production of hepatic glucose. (23). Normally, to balance the insulin resistance during pregnancy, the pancreatic β-cells increases insulin secretion (26). The mechanism behind the insulin resistance during pregnancy is yet to be revealed fully, but the elevated hormone and cytokine levels and their metabolic effect during pregnancy can be partially related. Potential hormones for the observed effect

are human placental lactogen (HPL), progesterone, prolactin and cortisol (23). Moreover, lots of hormones are produced by placenta to help the fetal development, some of which can block the function of insulin. The hormones such as cortisol and estrogen have strong diabetogenic effects (17), gradually leading to the insulin resistance.

During pregnancy, the fetal-placental unit development causes endocrine changes that trigger a shift in maternal nutrient metabolism. Human chorionic gonadotropin (hCG) is neutral in respect to glucose metabolism. Insulin binding is increased by estradiol. Human placental growth hormone, human placental lactogen (hPL) and progesterone, cortisol, induce insulin resistance, each on its own. Increase in the levels of above hormones can have cellular effects to cause the increasing resistance. (27). As pregnancy progresses and the placenta grows, hormone production also increases, and so does the level of IR. Insulin resistance become apparent between 20 and 24 weeks of pregnancy and continue to rise till the 35th week, when the growth of the placenta stops. The fact that IR rapidly abates following delivery suggests that placental hormones contribute to this state (27, 28).

The maintenance of normal glucose homeostasis during pregnancy depends on the capacity of the pancreatic β-cells to noticeably increase the secretion of insulin, thus compensating for the severe physiologic insulin resistance IR. Under the influence of various mediators, β-cells faces structural and functional changes that include β-cell mass increment and proliferation, increased level of insulin synthesis, and enhanced level of glucose-stimulated insulin secretion (29). Preclinical studies have suggested that both hPL and prolactin play a vital role in β-cell adaptation and increased insulin secretory capacity, which is critical for maintaining optimal glucose homeostasis during pregnancy. GDM arises in women with insufficient β-cell compensatory response, resulting in the hyperglycemia by which GDM is diagnosed. Thus, both pancreatic β-cell insufficiency and increase in insulin resistance level, can cause the GDM which is also affected by genetic predisposition and some other factors.

2.1.5 Genetic basis of GDM

Genetics may play a role in the progression of GDM (25, 30). Although it has long been known that this disease has a genetic basis (31) There have been few predisposing genes discovered with significant and reproducible effects. The genetics of GDM has been researched less than those of T2DM (32). However, there has been evidence of risk allele concordance as well as the direction of their effect. Only a few GDM genetic markers have been found in Bangladesh to yet (33, 34).

2.1.6 Risk factors of GDM

Some risk factors of GDM are maternal age, BMI > 30 kg/m2, family history of diabetes, history of previous gestational diabetes mellitus, history of abortion, history of macrosomia, history of gestational hypertension, history of preeclampsia, history of childbirth with congenital malformations, lipid levels increasing from the first trimester to the third trimester, ethnicity, FPG and TG levels in early pregnancy, and multigravida (8, 35-37). Furthermore, growing evidence suggests that hypertriglyceridemia and menarche are associated with gestational diabetes mellitus (38). Poor GDM knowledge, attitude, and practice can complicate the pregnancy and lead to negative outcomes. GDM not only causes immediate maternal complications like hypertensive disorders, hypothyroidism, need for cesarean section, IUGR, PROM, abortion, polyhydramnios, etc., and neonatal complications like hypoglycemia, respiratory distress, macrosomia, jaundice, large for gestational age, and stillbirth, but also increases the risk of future type 2 diabetes in the mother as well as the baby (39).

2.2 Type 2 diabetes mellitus (T2DM):

T2DM is a common but complex disease. To find out risk factors for diseases of this type is one of the central goals of human genetics. A number of analytical tools, different technologies and study designs are available for identifying these risk factors. In developed countries approximately 87%-91% diabetes cases are T2D and 7%-12% patients are estimated to have T1D (25). The second form of diabetes focused in this study is type 2 diabetes mellitus or T2DM. It is a polygenic metabolic disorder. Either

insulin resistance or reduced insulin secretion or combination of both results in increase blood glucose level. T2D is characterized by impaired insulin secretion from β cells coupled with insulin resistance in target tissues such as the liver, muscles and adipose tissue (40, 41). In addition, 1%-3% are estimated to have other types of diabetes (25).

2.2.1 Genetic pathophysiology of T2D

To date, almost 250 genetic variants have been identified that contribute to the risk of T2D (42). There are two main hypotheses on the genetic pathology of T2D; one is ‘common disease, common variant’ and another is ‘common disease, rare variant’ (43). In the ‘common disease, common variant’ theory, it is hypothesized that common variants (MAF >5%) with small effect size and low penetrance can cause the disease (43). On the other side, according to the ‘common disease, rare variant’ view, rare variants (MAF>1%) with large effect sizes and high penetrance might be the dominant cause of the disease (43). In previous genetic research, most of the T2D genetic studies focus on the standard variant. However, most of the identified susceptibility loci have petite effect sizes. They account for only a fraction of the apparent heritability, and most of them are located outside the coding regions (44, 45). Rare variants with more significant effects have been suggested to explain more of the ‘missing heritability; however, this has yet to be uncovered (46, 47). Furthermore, the inheritance model and risks of T2D differ across different ethnicities in a similar environment, which supports that there is a more complex genetic architecture underlying the pathology of T2D (48, 49)

2.3 Diabetes in pregnancy (DIP)

As opposed to GDM, diabetes in pregnancy (DIP) is defined as pregestational or preexisting diabetes (type 1 or type 2) and diagnosed when FPG is 7.0 mmol/L or 2-h PG is 11.1 mmol/L, according to 2013 WHO criteria (50, 51). The GDG updating the WHO recommendations acknowledged the distinction between DIP and GDM suggested by IADPSG, but proposed a slightly different language – "diabetes" rather than "overt diabetes" proposed by IADPSG. This distinction between diabetes and GDM is a novel proposal, and there is a paucity of published data on the consequences of applying it (50, 51).

2.4 GDM and the risk of diabetes

Women with GDM have a 17-63 % higher risk of T2DM 5-16 years after diagnosis (52). According to a study conducted in northeastern Ontario (53) 70% of women with GDM went on to develop T2DM, and the average time from GDM diagnosis to developing T2DM had been three years. T2DM was found in 6.9% of women with GDM after five years (95% CI: 3.8%-9.9%) and 21.1% of women with GDM after ten years (95% CI: 14.1%-27.5%) (54) . High postpartum body weight (>7 kg) increased the risk of diabetes by 86% and impaired glucose tolerance, impaired fasting glucose or pre-diabetes by 32%, but decreased postpartum body weight diminished the risk of pre-diabetes by 45% (55).

2.5 GDM and T2DM: share common pathophysiological Background

Like T2D, GDM also shows association with insulin resistance (IR) and inadequate compensatory secretion of insulin. During normal pregnancy, physiological IR does not result in dysglycaemia because of increased compensatory insulin secretion. GDM develops as a result of either abnormally high IR, maybe as a result of pre-existing IR in overweight women, or insufficient β-cell growth and subsequent insulin insufficiency (56). In GDM, pancreatic β-cell dysfunction occurs making insulin secretion difficult concerning glycemia and IR severity (57). Compared to the Westerners the pancreatic β-cell mass is relatively smaller and the insulin secretory capacity is also lower in Asians (58). In addition, in South Asian pregnant women, during pregnancy the β-cell adaptation was significantly lower in comparison to the Western Europeans (59). Women who have previous history of GDM and postpartum glucose tolerant record, they continue having high insulin resistance as well as β-cell dysfunction, on the other hand, non-GDM women showed remarkable decrease in IR. The persistence of impaired β—cell function in the presence of elevated IR raises their risk of developing diabetes in the future (60). For this reason, GDM is also known to be as an antecedent of type 2 diabetes.

2.6 Single-nucleotide polymorphisms (SNPs)

A single-nucleotide polymorphism (SNP) refers to variation in a DNA sequence that occurs when single nucleotides i.e. A, T, C, or G in the genome shows difference among the members of a biological species or in paired chromosomes of entities: Transitions (C/T or G/A) and transversions (C/G, A/T, C/A, or T/G) are the two separate categories of SNPs. In principle, SNPs at any particular site could be bi, tri, or tetra allelic. However, tri and tetra allelic SNPs are rare, and SNPs are generally biallelic in practice. SNPs are randomly distributed over the genome. In humans, SNPs are approximately 0.5% per nucleotide site (61). The bulk of the natural genetic variation in organisms is represented by SNPs or small insertions or deletions (62). Usually, SNPs are be fall in non-coding regions more frequently compared to the coding regions. They occur where natural selection acts and fix the allele of the SNP constituting the most favorable genetic adaptation (63). Recombination and mutation rates can determine the SNP density (64). SNPs within the coding region change the amino acid sequence of the synthesized protein, known as replacement polymorphism. It may be either missense, which results in different amino acids, or nonsense, which results in a premature stop codon. For example, a missense SNP (A-T) within the sixth codon of the β hemoglobin gene results in replacing the glutamic acid by valine, which causes Sickle cell anemia (65). Due to the degeneracy of the genetic code, sometimes SNPs located in the coding sequence have no effect on the protein amino acid sequence, hence are called synonymous polymorphism. Among different types of genetic variants, SNPs were targeted in this study. SNPs influence disease risk, drug efficacy, and side effects. SNPs can be divided into linked SNP and causative SNPs. Linked SNPs are located outside the gene and do not affect protein production or function. Causative SNPs located inside the gene change protein production, structure, and function.

2.7 CDK5 Regulatory Subunit Associated Protein 1-Like1(*CDKAL1)*

CDKAL1 is a marker of insulin secretion impairment that raises the risk of T2DM. The relevance of the CDKAL1 gene in pancreatic β-cell function is unknown. According to a mouse study, *CDKAL1* knockout animals had decreased conversion of proinsulin to insulin and lower ATP synthesis in mitochondria after glucose stimulation. (66). The

rs7756992 SNP is found on 6p22.3 in intron 5 of the CDK5 regulatory subunit associated protein 1-like 1 (CDKAL1) gene. It is found in a 201.7-kb LD block that contains exons 1–5 of the CDKAL1 gene as well as the minimum promoter region, but no other known genes. In the Chinese Han population, a genetic study discovered a relationship between T2DM and *CDKAL1* SNPs at rs10946398, but not at rs736425 or rs4712527. (67). In Asian, Caucasian, African, and Arab groups, the meta-analysis study (68) found a substantial connection between T2DM and *CDKAL1* variants at rs7754840 and rs7756992, similar to the findings in Japanese and Lebanese populations (69-73). The meta-analysis study also found a relationship between rs10916398 and diabetes in Asian, Caucasian, and African populations (68). Furthermore, the *CDKAL1* variant (rs10916398) was found to increase the risk of T2D in a Caucasian population (74, 75). In East Asian and European populations, additional SNPs (rs4712524, rs9295475, and rs9460546) linked to T2D were discovered (76). The association between T2D and rs2237892 was discovered in a GWAS of a Japanese population (77). A GWAS of a Caucasian sample also discovered that the rs7754840 variant increased the risk of T2DM (78). GDM risk is increased by *CDKAL1* polymorphisms at rs7756992 and rs7754840 (79). GWAS in a Korean population validated the association between the SNP in CDKAL1 (rs7754840) and GDM (80). However, a study on the Chinese population found no link between the SNP rs7754840 and GDM (81). In addition, the study in Danish populations revealed correlation of rs7756992 with GDM (82). The GWA study found a link between CDKAL1 (rs2206734) and BMI in a Japanese population (83). The polymorphism at rs10946398 was shown to be unrelated to BMI in a Chinese population (84). GWASs suggest that the CDKAL1 risk allele rs7754840 is linked to a reduction in insulin secretion (78) and a 24% reduction in first-phase insulin release, a larger glucose area under the curve, and insulin release impairment (85).

2.8 Fat mass and obesity associated gene (*FTO)*

The *FTO* gene is located on chromosome 16 (16q12.2), containing nine exons and several SNPs (86). Some studies have revealed no relationship between FTO polymorphisms and FTO expression or function (87), while others have claimed that FTO variants play an important role in controlling body weight and fat mass via modulating food intake (88). SNPs in *FTO* have been found to influence obesity

through modifying the expression of the neighboring genes IRX3 and RPGRIP1L (89). Although the mechanisms affecting T2DM caused by these noncoding polymorphisms are unknown, variants in *FTO* can build long-range functional connections with IRX3, a determinant of body mass and composition (90). Furthermore, recent research suggests that hepatic FTO contributes to glucose homeostasis (91-93), implying that *FTO* may be involved in carbohydrate metabolism regulation. Intron 1 of the *FTO* gene contains rs8050136, an area of significant linkage disequilibrium (94). Although recent study has discovered a link between *FTO* SNPs and the risk of GDM, other studies have found the opposite, therefore no clear conclusion has been established (95-97). At least in the Korean population, the rs8050136 does not raise the incidence of GDM but may provide protection by improving insulin secretory ability. As observed in the recent study, *FTO rs8050136* may influence insulin indices directly or indirectly. A probable explanation is that control persons with the FTO risk gene are predisposed to insulin resistance and can boost insulin secretion to compensate for inadequate insulin sensitivity in order to maintain perfect glucose homeostasis. As a result, the pathophysiology of FTO's role to T2D should be clarified.

FTO had a sequence in common with iron- and 2-oxoglutarate-dependent oxygenases, and the quantity of *FTO* mRNA in the hypothalamus was affected by feeding and fasting (98). Postnatal development retardation (shorter body length, lower body weight and relatively poor bone mineral density) and decreased insulin-like growth factor 1 (IGF-1) levels were seen in mice with the *FTO* mutation. (99). The AT-rich interactive domain 5B (MRF1-like) (ARID5B)-mediated regulation of Iroquois homeobox 3 (IRX3) and Iroquois homeobox 5 (IRX5) is disrupted by the FTO polymorphisms (IRX5). IRX3 and IRX5 inhibition causes a cell-autonomous switch from white adipocyte browning to mitochondrial thermogenesis, leading in increased fat storage and body weight (100). In European populations, variation in *FTO* (rs8050136) was linked to the risk of T2DM (74, 75). In a Lebanese Arab population, a link between T2DM and FTO polymorphisms (rs8050136 and rs17817449, except rs1121980) was observed. (101). After adjusting for age, sex, and BMI, the meta-analysis study based on European and East Asian populations (102) discovered a relationship between T2DM and polymorphism in FTO at rs9939609, which is similar to the studies in Norwegian and Swedish populations (103). Furthermore, in a study in

Spain, rs9939609 enhanced the incidence of GDM (104). Additional SNPs linked to T2DM were identified in a Chinese population, including rs6499640 and rs3751812 (105). However, genetic research found no association between rs9939609 and diabetes risk in a Japanese group (69), as well as rs8050136 in African American and Chinese populations (106, 107). In the Danish research, there was no association reported between rs9939609 and GDM (82). Obesity and rs9939609 in white Americans and rs1421085 in African Americans were found to be linked in a large prospective study in the United States. (108). A meta-analysis demonstrated that the rs9939609 of *FTO* gene increased the number of overweight and obese people. (88, 109, 110). A GWA study identified correlation between obesity and a number of SNPs (rs9930506, rs8050136, rs1121980, rs7193144, rs9939609, rs9926289, rs6602024, rs7907949, rs965670, rs1188445, and rs6965526) of the *FTO* gene (111). A subsequent GWAS in a European population identified another FTO variant (rs1421085) linked to obesity (112). BMI and T2DM have both been linked to common FTO gene variants; the tendency to T2DM can be entirely explained by the weight-increasing effect (88). In agreement with other studies, it was observed that *FTO* predisposes to metabolic syndrome primarily through its obesity-related effects. Although the processes by which *FTO* variants enhance the risk of obesity are unknown, because FTO is highly expressed in the hypothalamus, they could involve impacts on appetite regulation. (113).

2.9 Heat Shock Protein Family A (Hsp70) Member 1 Like *(HSPA1L)*

This substitution *HSP70-hom* +2437 C/T polymorphism (i.e, Thr→Met) amino acid substitution at position 493) may be associated with variations in the peptide-binding specificity of different *HSP70-hom* haplotypes (32). The 70-kDa heat shock protein (HSP70) family is the most abundant in eukaryotic cells and is essential for cell survival under stressful conditions (114). Synthesis of these highly conserved molecular chaperones was induced by stresses such as heat shock, ischemia, and other types of cellular stresses (115, 116). Three main genes (*HSPA1A*, *HSPA1B*, and *HSPA1L*) of HSP70 family are observed in human for which the coding proteins are called as HSP70-1, HSP70-2, and HSP70-hom, respectively. HSPs are also involved in diabetes by creating effect on insulin sensitivity (117). The gene variants found in *HSP70* showed association with the increased risk of T2DM development (32, 118-122). The

location + 2437 of HSP70-homologous is one of highly studied SNPs of HSPA1L. This is a nonsynonymous missense mutation in HSP70-hom that results in a shift in amino acid from threonine to methionine, which alters the protein's stability and activity (32, 123). Furthermore, this SNP may have an impact on HSP70 expression or function, contributing to disease susceptibility and stress tolerance. (124). A number of epidemiological studies have been conducted to determine the link between this polymorphism and obesity, diabetes, and diabetes nephropathy (119). The HSP70-hom + 2437 T/C polymorphism was found to have a substantial and nonsignificant correlation with T2DM in these investigations (32, 119).

2.10 Peroxisome proliferator-activated receptor gamma (*PPARG*)

Peroxisome proliferator-activated receptors (PPARs) are transcription factors that are ligand activated and belongs to the nuclear hormone receptor superfamily (125). Separate genes encode three isoforms: PPAR-γ, PPAR-α, and PPAR-β/ δ, which share 60% to 80% homology in their ligand- and DNA-binding domains, as well as they exhibit distinct patterns of expression. They also show overlapping as well as distinct biological activities (126). *PPARG* have significant role in glucose homeostasis which is molecular target of thiazolidinediones (TZDs, a class of insulin-sensitizing drugs). Thiazolidinediones are PPARGZ ligands widely used for the treatment of type 2 diabetes (125). Insulin action, adipocyte development, lipid storage, and fat-specific metabolism are all associated with this gene (127). PPARG improves glucose homeostasis by activating glucose transporter 2 and glucokinase in the liver and pancreatic β-cells (128). PPARG also increases insulin sensitivity in peripheral tissue as well as glucose sensitivity in liver and pancreatic β-cells. In comparison to wild-type mice, PPARG knockout ones had a stronger insulin-induced increase in glucose disposal rate and a greater insulin-induced inhibition of hepatic glucose synthesis (129). PPARG gene was associated with T2DM by Candidate Gene Association Studies (CGAS) and has been verified in multiple studies (78, 130, 131). The CGAS in Chinese (106), Japanese (132), and Indian (133) populations and the GWAS in Finnish (74) and Caucasian people (78) found that rs1801282 raised the risk of T2DM. However, no link was found between T2DM and *PPARG* variants (rs1801282, rs12636454, and rs11128597) in a genetic investigation in the Chinese Han population (67). A French

study revealed a link between GDM and PPARG variants (rs1801282 and rs3856806). (134). However, a candidate gene approach used in Sweden, Denmark, and Korea found no link between PPARG variant (rs3856806 and rs1801282) and cancer (79, 82, 135). In menopausal women, a link between rs1801282 and total body fat mass has been proposed (136).

2.11 Transcription factor 7 like 2 (*TCF7L2*)

The *TCF7L2* gene encodes a high mobility group (HMG) box-containing transcription factor involved in blood glucose homeostasis. TCF7L2 controlled proglucagon in enteroendocrine cells by inhibiting the glucagon gene through the Wnt signaling pathway (137). *TCF7L2* is involved in pancreatic cell formation as well as glucose-induced insulin secretion (138). Decreased TCF7L2 protein levels was found in the pancreas of T2DM patients compared with healthy individuals. TCF7L2 and glucagon-like peptide-1 (GLP1R) and glucose-dependent insulinotropic polypeptide (GIPR) interactions may influence pancreatic-cell activity and survival, according to this study (139). The first GWAS paper reported the discovery of four novel T2D loci, including the *TCF7L2*, in 2007 (140). Most of them have been found in recent years based on meta-analyses (141). The meta-analysis of GWAS data has proved that large sample sizes are necessary to identify the minor effects of susceptible SNPs (42). The risk *TCF7L2* variant is still regarded as the most influential common T2D variant (OR=1.46) (142). Three of the selected single nucleotide polymorphisms (SNPs) have previous record of identification in candidate gene or linkage studies, which showed reproducible results in other studies. (131, 143, 144). Persons with IGT showed increased risk of diabetes when they possess the rs7903146 in intron 3 of the *TCF7L2* gene confirming significant association(144). The rs7903146 raises the risk of T2DM in Dutch, Han Chinese, British, Korean, Chinese, African American, Arabic, and Indian populations (72, 75, 107, 133, 145-147) and the risk of GDM in Scandinavian (135), Korean (79), Danish (82), and Czech (148) women. The Chinese study (67) also showed a relationship between T2DM and two SNPs (rs7903146 and rs6585205). In a North Indian population, an association between rs10885409 and T2D was observed (149). However, a Dutch investigation found no association between the TCF7L2 gene variant rs4430796 and T2DM (150). The GWAS in the French, Finnish and Caucasian

populations confirmed the association between T2D and rs7903146 (74, 78, 140). Furthermore, the Japanese GWAS found that rs7901695 increased the chances of T2D (77). Another SNP (rs114748339) linked to T2D was discovered in a meta-analysis of GWAS in African Americans(151). Additional SNPs (rs12255372, rs4506565, and rs7901695) associated to GDM were discovered in candidate gene investigations in Austria (152), Spain (104), and the Czech Republic (148) However, in Korea (79) and Denmark rs12255372 in *TCF7L2* did not correlate with GDM.

2.12 Wolfram syndrome 1 (*WFS1*)

The WFS1 gene produces wolframin, a transmembrane glycoprotein that helps the endoplasmic reticulum maintain calcium homeostasis. Wolfram syndrome (OMIM 222300) is caused by mutations in this gene and is characterized by diabetes insipidus, juvenile-onset non-autoimmune diabetes mellitus, optic atrophy, and deafness (153). Depending on the genetic background of the mice, disruption of the WFS1 gene induces overt diabetes or decreased glucose tolerance. Wolframin deficiency was found in both people and animals, and both demonstrated a decrease of pancreatic beta cells. (154). WFS1 is thus required for the survival and function of insulin-producing beta cells in the pancreas. In Caucasians, the rs10010131 of the *WFS1* locus has been demonstrated to be related with T2DM (155). The association between variants in *WFS1* and risk of type 2 diabetes was replicated (156) in the European population. Significant association was observed among the major alleles of the 3 variants of *WFS1* rs10010131, rs1801213/ rs7672995, and rs734312 and prevalent T2D in the DESIR cohort (157). Moreover, in the haplotype block containing the*WFS1* gene, the most frequent haplotype showed involvement in modulation of insulin secretion and showed association with increased risk of T2D. In a latest meta-analysis, rs10010131 of the *WFS1* gene has been related to the risk of T2DM (156). The rs10010131 is located in intron 4 and is in high LD with the other high-risk variants of this gene (146, 155, 156, 158, 159) . *WFS1*-encoded protein has recently been described as a component of the unfolded protein response with a critical role in maintaining endoplasmic reticulum homeostasis in pancreatic β-cells (160). As a result, SNPs in the WFS1 gene may affect beta-cell activity and GLP-1 responsiveness by disrupting endoplasmic reticulum homeostasis. An impaired or malfunctioning GLP-1 impact could lead to decreased

postprandial insulin secretion, as well as influence β-cell proliferation and differentiation stimulation (161). WFS1 is a 890-amino-acid transmembrane polypeptide that is widely expressed, with high levels of expression in pancreatic islets and particular neurons, and subcellular localization to the endoplasmic reticulum (162). Due to the general clinical symptoms of diabetes insipidus, young-onset non-immune insulin-dependent diabetes mellitus, optic atrophy, and deafness, WFS is also known as the DIDMOAD syndrome. (163). Mice missing Wfs1 demonstrated increasing β-cell loss and insulin secretion impairment (154). Endoplasmic reticulum stress and apoptosis resulted in decreased β-cell survival (164, 165).

2.13 Case-control Studies

Case-control study is a form of observational research. The investigator does not change the exposure status in an observational research (166). The researcher examines the relationship between exposure and result in study participants. Participants in a case-control study are chosen for the study depending on their outcome status. As a result, some individuals have the desired outcome (referred to as cases), while others do not (referred to as controls). After that, the investigator evaluates the exposure in both groups. As a result, in a case-control study, the outcome must occur in at least some of the individuals. As shown in Figure 2.2, some research participants had the outcome (cases) at the time of enrollment into the study (sampling of participants), while others do not (controls). The exposure of interest will be assessed in both cases and controls during the study methods. The relationship between exposure and outcome in these research participants will subsequently be investigated (167).

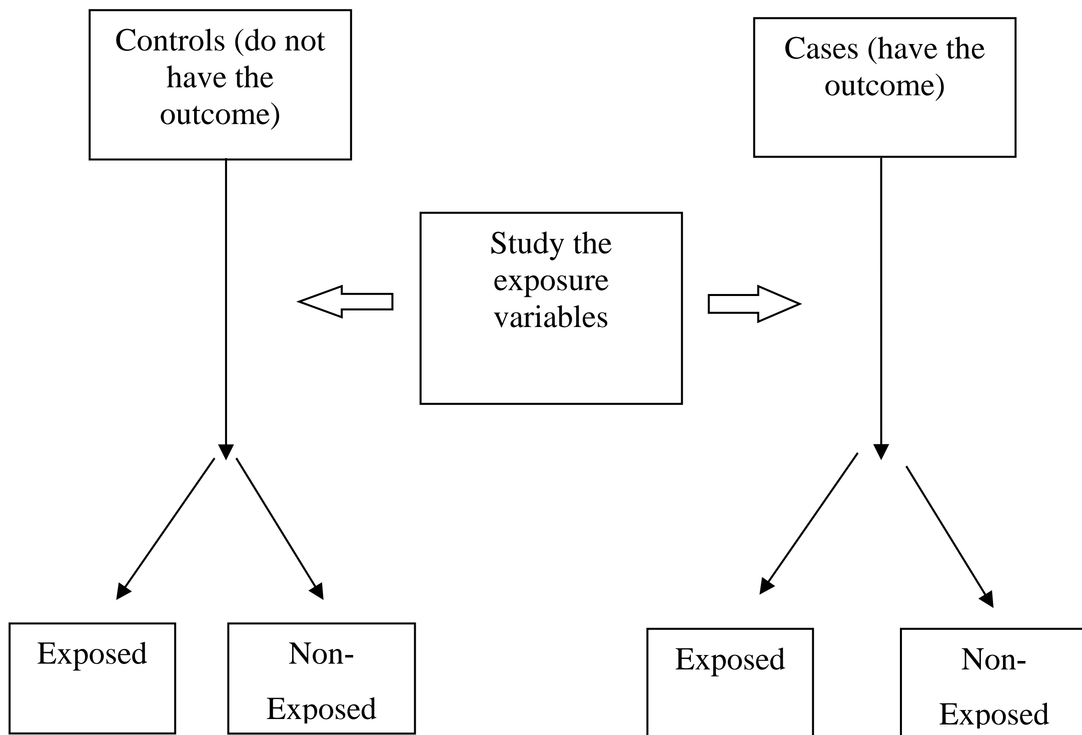


Figure 2.2: Schematic example of Case control Study

2.14 Odds ratio

The cross product ratio is the ratio of the two pseudo-rates in a case control study, which is commonly expressed as A1B0/A0B1. In a case-control study, the cross product ratio is calculated by dividing the ratio of cases to controls among exposed participants (A1/B1) by the ratio of cases to controls among unexposed patients (A0/B0). The exposure odd ratio can also be defined as the probability of being exposed among cases (A1/A0) divided by the probability of being exposed among controls (B1/B0). While both interpretations yield the same result, viewing this odds ratio as a ratio of case-control ratios demonstrates more clearly how the control group serves as a denominator in a cohort study and how the ratio of pseudo frequencies yields the same result as the ratio of incidence rates, incidence proportion, or incidence odds in the source population if sampling is exposure-independent. In a case-control research, the odds

ratio is the "measure of association." In a case-control study, it quantifies the connection between an exposure (such as consuming a food or attending an event) and a disease. The odds ratio is computed by taking the number of case-patients who did or did not have exposure to a factor (such as a certain diet) and dividing it by the number of controls who did or did not have exposure. The odds ratio indicates how much more likely case-patients are to be exposed than controls. An odds ratio of

• A value of **1.0 (or close to 1.0)** suggests that the chances of case-patients being exposed are the same as, or similar to, the chances of controls being exposed. The sickness is not linked to the exposure.
• A value of **more than 1.0** suggests that case-patients have a higher chance of being exposed than controls. The sickness could be worse by the exposure.
• A value of **less than 1.0** suggests that case-patients have a lower chance of being exposed than controls. It's possible that the exposure will safeguard one from getting sick.

2.15 Candidate polymorphism studies

Candidate polymorphism studies are investigations of genotype–trait relationships for which an a priori hypothesis about functioning exists. A genetic variant at a single site inside a gene is referred to as a polymorphism. In order to be categorized as a polymorphism, a variation must be found in at least 1% of a population. SNP stands for single nucleotide polymorphism. Prior scientific evidence indicating that the set of polymorphisms under inquiry is relevant to the illness trait is often used in candidate polymorphism investigations. The fundamental premise is that the variable location under research is functional, and the goal is to test for the presence of association. Candidate polymorphism studies seek to identify whether a single SNP or a group of SNPs has a direct impact on a disease trait.

2.16 HOMA indices in GDM

HOMA indices are commonly used in GDM studies as a measures of insulin resistance and sensitivity. In a cross-sectional study, significantly higher HOMA-IR was observed among GDM than NGT but HOMA-B values were similar in both groups (168).

However, both HOMA-IR and HOMA-B were higher in GDM when compared with non-pregnant healthy controls. Higher HOMA-IR has been reported in GDM than NGT which remained high in six months postpartum follow up (169). Another study which compared HOMA-IR in obese and non-obese pregnant women, demonstrated elevated HOMA-IR in obese than non-obese (170). South Asians had higher HOMA-IR compared with Western Europeans, at early gestation and at 24 weeks (171). Even the increment of HOMA-B from early to later part of gestation was less for the South Asian than the Western European women. South Asians had higher HOMA-B when compared with Western Europeans at early gestation but not at 24 weeks(171). There was no significant difference was observed in HOMA-IR and HOMA-S% among lean Bangladeshi individuals with GDM and NGT whereas HOMA-B is significantly lower in GDM than NGT indicating that insulin secretory defect may be the major determinant of GDM in lean mothers (172). After evaluating HOMA indices in pregnant women in our population, higher HOMA-IR was observed in GDM than NGT where HOMA-B was low among GDM and so as HOMA-S%(173).

2.17 Hardy Weinberg Equilibrium (HWE)

The Hardy-Weinberg (HW) principle refers to the connection between allele frequencies and counts of genotype in successive generations in the absence of any disturbing factors. It predicts that genotype and allele frequencies should remain unchanged from generation to generation in an infinitely large random mating population (174). The Hardy-Weinberg equilibrium (HWE) is a non-evolutionary model that reflects one of the most fundamental ideas in population genetics and evolutionary biology (175). One of the most significant ideas in population genetics, the HW principle, was developed to examine allele frequency variations in a population over generations. (176). It is currently commonly used to identify inbreeding, population stratification, and genotyping errors in human illness research. The asymptotic Pearson's chi-square goodness-of-fit test and the exact test are the most prevalent ways for assessing departure from the Hardy–Weinberg proportions in data. Although the Pearson's chi-square goodness-of-fit test is basic and easy, it is extremely sensitive to small sample sizes or unusual allele frequencies. In these cases, an accurate test of HW proportions is recommended. Complete enumeration of

heterozygote genotypes or the Markov chain Monte Carlo technique can be used to perform the exact test. Methods to identify genotyping error have been developed since no genotyping method is 100% reliable, and genotype errors can lead to increased random error and bias in gene-disease associations (177). HWE tests are commonly used to check genotype information quickly. (176, 178-181). Population substructure, purifying selection, copy number variation, or genotyping error are all possible reasons for HWE departure. (182).

2.18 Confounder

Comparing the estimated measure of association before and after controlling for confounding is a simple and direct technique to determine if a given risk factor produced confounding. To put it another way, compute the measure of association before and after controlling for a possible confounding factor. Confounding was observed if the discrepancy between the two measures of association was 10% or more (183). If the percentage is less than 10%, there was little, if any, confounding. The specifics of how to do this will be discussed further down. Other researchers will look into whether a potential confounding variable is linked to the exposure and whether it is linked to the outcome of interest. The variable is considered a confounder if there is a clinically relevant association between it and the risk factor, as well as between it and the result (regardless of whether that relationship reaches statistical significance). Other researchers do formal hypothesis tests to see whether a variable is linked to the exposure of interest and the outcome (184).

References

1. International Diabetes Federation. Gestational diabetes 2020 [updated 07/01/2020. Available from: https://idf.org/our-activities/care-prevention/gdm.html.

2. Noctor E, Dunne FPJWjod. Type 2 diabetes after gestational diabetes: the influence of changing diagnostic criteria. 2015;6(2):234.

3. Miller HC. The effect of diabetic and prediabetic pregnancies on the fetus and newborn infant. The Journal of pediatrics. 1946;29(4):455-61.

4. Carrington ER, Shuman CR, Reardon HS. Evaluation of the prediabetic state during pregnancy. Obstetrics and gynecology. 1957;9(6):664-9.

5. O'Sullivan JB, Mahan CM. CRITERIA FOR THE ORAL GLUCOSE TOLERANCE TEST IN PREGNANCY. Diabetes. 1964;13:278-85.

6. American Diabetes Association. Gestational diabetes mellitus. Diabetes care. 2003;26 Suppl 1:S103-5.

7. Metzger BE, Coustan DR. Summary and recommendations of the Fourth International Workshop-Conference on Gestational Diabetes Mellitus. The Organizing Committee. Diabetes care. 1998;21 Suppl 2:B161-7.

8. Lee KW, Ching SM, Ramachandran V, Yee A, Hoo FK, Chia YC, et al. Prevalence and risk factors of gestational diabetes mellitus in Asia: a systematic review and meta-analysis. 2018;18(1):1-20.

9. Sandesh-Panthi MA, Hasanat M-H, Yasmin-Aktar N-S, Sharmin-Jahan MFJJ. Frequency of gestational diabetes mellitus in Bangladesh impact of WHO 2013 screening criteria: Efficiency DIPSI & WHO 1999 criteria. 2015;2(2):13-9.

10. Shaat N, Groop L. Genetics of gestational diabetes mellitus. Current medicinal chemistry. 2007;14(5):569-83.

11. Hadden DR. Diabetes in pregnancy 1985. Diabetologia. 1986;29(1):1-9.

12. Nusrat-Sultana HM, Sharmin-Jahan M-H, Sandesh-Panthi Y-A, Fariduddin MJStJ. Alarming frequency of Gestational Diabetes Mellitus (GDM) attending a tertiary care hospital in Bangladesh. 2016.

13. Wahi P, Dogra V, Jandial K, Bhagat R, Gupta R, Gupta S, et al. Prevalence of gestational diabetes mellitus (GDM) and its outcomes in Jammu region. 2011;59(4):227-30.

14. Alfadhli EMJSmj. Gestational diabetes mellitus. 2015;36(4):399.

15. Nguyen CL, Pham NM, Binns CW, Duong DV, Lee AHJJodr. Prevalence of gestational diabetes mellitus in eastern and southeastern Asia: a systematic review and meta-analysis. 2018;2018.

16. Sermer M, Naylor C, Gare D, Kenshole A, Ritchie J, Farine D, et al. for the Toronto Tri-Hospital Gestational Diabetes Investigators: Impact of increasing carbohydrate intolerance on maternal-fetal outcomes in 3,637 women without gestational diabetes. 1995;173(1):146-56.

17. Carr DB, Gabbe SJCD. Gestational diabetes: detection, management, and implications. 1998;16(1):4-12.

18. Kjos SL, Buchanan TAJNEjom. Gestational diabetes mellitus. 1999;341(23):1749-56.
19. Landon MB, Galan HL, Jauniaux ER, Driscoll DA, Berghella V, Grobman WA, et al. Obstetrics: Normal and Problem Pregnancies E-Book: Elsevier Health Sciences; 2020.
20. Silverman BL, Metzger BE, Cho NH, Loeb CAJDc. Impaired glucose tolerance in adolescent offspring of diabetic mothers: relationship to fetal hyperinsulinism. 1995;18(5):611-7.
21. Seshadri RJDC. American diabetes association gestational diabetes mellitus. 2002;25:S94-S6.
22. Silverman BL, Rizzo T, Green OC, Cho NH, Winter RJ, Ogata ES, et al. Long-term prospective evaluation of offspring of diabetic mothers. 1991;40(Supplement_2):121-5.
23. Lain KY, Catalano PM. Metabolic changes in pregnancy. Clinical obstetrics and gynecology. 2007;50(4):938-48.
24. Catalano PM, Tyzbir ED, Wolfe RR, Roman NM, Amini SB, Sims EA. Longitudinal changes in basal hepatic glucose production and suppression during insulin infusion in normal pregnant women. American journal of obstetrics and gynecology. 1992;167(4 Pt 1):913-9.
25. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes research and clinical practice. 2018;138:271-81.
26. Buchanan TA, Xiang A, Kjos SL, Watanabe R. What is gestational diabetes? Diabetes care. 2007;30 Suppl 2:S105-11.
27. Ramos-Román MA. Prolactin and lactation as modifiers of diabetes risk in gestational diabetes. Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme. 2011;43(9):593-600.
28. Buchanan TA, Xiang AH. Gestational diabetes mellitus. The Journal of clinical investigation. 2005;115(3):485-91.
29. Rieck S, Kaestner KH. Expansion of beta-cell mass in response to pregnancy. Trends in endocrinology and metabolism: TEM. 2010;21(3):151-8.

30. Gjesing AP, Rui G, Lauenborg J, Have CT, Hollensted M, Andersson E, et al. High Prevalence of Diabetes-Predisposing Variants in MODY Genes Among Danish Women With Gestational Diabetes Mellitus. Journal of the Endocrine Society. 2017;1(6):681-90.

31. Lambrinoudaki I, Vlachou SA, Creatsas G. Genetics in gestational diabetes mellitus: association with incidence, severity, pregnancy outcome and response to treatment. Current diabetes reviews. 2010;6(6):393-9.

32. Moniruzzaman M, Ahmed I, Huq S, All Mahmud MS, Begum S, Mahzabin Amin U, et al. Association of polymorphism in heat shock protein 70 genes with type 2 diabetes in Bangladeshi population. 2020;8(2):e1073.

33. Hasan M, Hasnat MA, Hasan KN, Panthi S, A Y, Sultana N, et al. TCF7L2 gene rs7903146 polymorphism is observed in gestational diabetes mellitus in Bangladesh. Integrative Obesity and Diabetes. 2016.

34. Amin USM, Parvez N, Rahman TA, Hasan MR, Das KC, Jahan S, et al. CDKAL1 gene rs7756992 A/G and rs7754840 G/C polymorphisms are associated with gestational diabetes mellitus in a sample of Bangladeshi population: implication for future T2DM prophylaxis. Diabetol Metab Syndr. 2022;14(1):18.

35. Larrabure-Torrealva GT, Martinez S, Luque-Fernandez MA, Sanchez SE, Mascaro PA, Ingar H, et al. Prevalence and risk factors of gestational diabetes mellitus: findings from a universal screening feasibility program in Lima, Peru. 2018;18(1):1-9.

36. Amiri FN, Faramarzi M, Bakhtiari A, Omidvar S. Risk Factors for Gestational Diabetes Mellitus: A Case-Control Study. Am J Lifestyle Med. 2018;15(2):184-90.

37. Rajput M, Bairwa M, Rajput R. Prevalence of gestational diabetes mellitus in rural Haryana: A community-based study. Indian J Endocrinol Metab. 2014;18(3):350-4.

38. Xiao Y, Chen R, Chen M, Luo A, Chen D, Liang Q, et al. Age at menarche and risks of gestational diabetes mellitus: a meta-analysis of prospective studies. Oncotarget. 2017;9(24):17133-40.

39. Goldman M, Kitzmiller JL, Abrams B, Cowan RM, Laros RK, Jr. Obstetric complications with GDM. Effects of maternal weight. Diabetes. 1991;40 Suppl 2:79-82.

40. Szoke E, Gerich JE. Role of impaired insulin secretion and insulin resistance in the pathogenesis of type 2 diabetes mellitus. Comprehensive therapy. 2005;31(2):106-12.
41. Cantley J, Ashcroft FM. Q&A: insulin secretion and type 2 diabetes: why do β-cells fail? BMC biology. 2015;13:33.
42. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nature genetics. 2018;50(11):1505-13.
43. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. Current opinion in genetics & development. 2009;19(3):212-9.
44. Langenberg C, Lotta LA. Genomic insights into the causes of type 2 diabetes. Lancet (London, England). 2018;391(10138):2463-74.
45. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.
46. Chen G, Zhang Z, Adebamowo SN, Liu G, Adeyemo A, Zhou Y, et al. Common and rare exonic MUC5B variants associated with type 2 diabetes in Han Chinese. PloS one. 2017;12(3):e0173784.
47. Majithia AR, Flannick J, Shahinian P, Guo M, Bray MA, Fontanillas P, et al. Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(36):13127-32.
48. van Valkengoed IGM, Argmann C, Ghauharali-van der Vlugt K, Aerts J, Brewster LM, Peters RJG, et al. Ethnic differences in metabolite signatures and type 2 diabetes: a nested case-control analysis among people of South Asian, African and European origin. Nutrition & diabetes. 2017;7(12):300.
49. Admiraal WM, Holleman F, Snijder MB, Peters RJ, Brewster LM, Hoekstra JB, et al. Ethnic disparities in the association of impaired fasting glucose with the 10-year cumulative incidence of type 2 diabetes. Diabetes research and clinical practice. 2014;103(1):127-32.
50. World Health Organization. Diagnostic Criteria and Classificationof Hyperglycaemia First Detected in Pregnancy.; 2013.

51. Panthi S, Hasanat M, Hasan M, Aktar Y, Sultana N, Jahan S, et al. Frequency of Gestational Diabetes Mellitus in Bangladesh: Impact of WHO 2013 Screening Criteria : Efficiency of DIPSI and WHO 1999 Criteria JOURNAL OF CLINICAL DIABETOLOGY. Journal of Clinical Diabetology. 2015;2:13-20.

52. Hanna FW, Peters JR. Screening for gestational diabetes; past, present and future. Diabetic medicine : a journal of the British Diabetic Association. 2002;19(5):351-8.

53. Mohamed N, Dooley J. Gestational diabetes and subsequent development of NIDDM in aboriginal women of northwestern Ontario. International journal of circumpolar health. 1998;57 Suppl 1:355-8.

54. Sivaraman SC, Vinnamala S, Jenkins D. Gestational diabetes and future risk of diabetes. Journal of clinical medicine research. 2013;5(2):92-6.

55. Ben-Haroush A, Yogev Y, Hod M. Epidemiology of gestational diabetes mellitus and its association with Type 2 diabetes. Diabetic medicine : a journal of the British Diabetic Association. 2004;21(2):103-13.

56. Park S, Kim MY, Baik SH, Woo JT, Kwon YJ, Daily JW, et al. Gestational diabetes is associated with high energy and saturated fat intakes and with low plasma visfatin and adiponectin levels independent of prepregnancy BMI. European Journal of Clinical Nutrition. 2013;67(2):196-201.

57. Metzger BE, Buchanan TA, Coustan DR, de Leiva A, Dunger DB, Hadden DR, et al. Summary and recommendations of the Fifth International Workshop-Conference on Gestational Diabetes Mellitus. Diabetes care. 2007;30 Suppl 2:S251-60.

58. Yang SJ, Kim TN, Baik SH, Kim TS, Lee KW, Nam M, et al. Insulin secretion and insulin resistance in Korean women with gestational diabetes mellitus and impaired glucose tolerance. 2013;28(3):306.

59. Mørkrid K, Jenum AK, Sletner L, Vårdal MH, Waage CW, Nakstad B, et al. Failure to increase insulin secretory capacity during pregnancy-induced insulin resistance is associated with ethnicity and gestational diabetes. 2012.

60. Cerf ME. Beta cell dysfunction and insulin resistance. Front Endocrinol (Lausanne). 2013;4:37-.

61. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

62. Mills RE, Pittard WS, Mullaney JM, Farooq U, Creasy TH, Mahurkar AA, et al. Natural genetic variation caused by small insertions and deletions in the human genome. Genome research. 2011;21(6):830-9.

63. Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L. Natural selection has driven population differentiation in modern humans. Nature genetics. 2008;40(3):340-5.

64. Sainudiin R, Clark AG, Durrett RT. Simple models of genomic variation in human SNP density. BMC genomics. 2007;8:146.

65. Ingram VM. Gene Mutations in Human Hæmoglobin: the Chemical Difference Between Normal and Sickle Cell Hæmoglobin. Nature. 1957;180(4581):326-8.

66. Wei FY, Suzuki T, Watanabe S, Kimura S, Kaitsuka T, Fujimura A, et al. Deficit of tRNA(Lys) modification by Cdkal1 causes the development of type 2 diabetes in mice. The Journal of clinical investigation. 2011;121(9):3598-608.

67. Lin Y, Li P, Cai L, Zhang B, Tang X, Zhang X, et al. Association study of genetic variants in eight genes/loci with type 2 diabetes in a Han Chinese population. BMC Medical Genetics. 2010;11(1):97.

68. Peng F, Hu D, Gu C, Li X, Li Y, Jia N, et al. The relationship between five widely-evaluated variants in CDKN2A/B and CDKAL1 genes and the risk of type 2 diabetes: a meta-analysis. Gene. 2013;531(2):435-43.

69. Horikawa Y, Miyake K, Yasuda K, Enya M, Hirota Y, Yamagata K, et al. Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. The Journal of clinical endocrinology and metabolism. 2008;93(8):3136-41.

70. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nature genetics. 2007;39(6):770-5.

71. Nemr R, Almawi AW, Echtay A, Sater MS, Daher HS, Almawi WY. Replication study of common variants in CDKAL1 and CDKN2A/2B genes associated with type 2 diabetes in Lebanese Arab population. Diabetes research and clinical practice. 2012;95(2):e37-40.

72. Ng MC, Park KS, Oh B, Tam CH, Cho YM, Shin HD, et al. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. Diabetes. 2008;57(8):2226-33.

73. Tan JT, Ng DP, Nurbaya S, Ye S, Lim XL, Leong H, et al. Polymorphisms identified through genome-wide association studies and their associations with type 2 diabetes in Chinese, Malays, and Asian-Indians in Singapore. The Journal of clinical endocrinology and metabolism. 2010;95(1):390-7.

74. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007;316(5829):1341-5.

75. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007;316(5829):1336-41.

76. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, et al. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nature genetics. 2008;40(9):1098-102.

77. Yamauchi T, Hara K, Maeda S, Yasuda K, Takahashi A, Horikoshi M, et al. A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at UBE2E2 and C2CD4A-C2CD4B. Nature genetics. 2010;42(10):864-8.

78. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007;316(5829):1331-6.

79. Cho YM, Kim TH, Lim S, Choi SH, Shin HD, Lee HK, et al. Type 2 diabetes-associated genetic variants discovered in the recent genome-wide association studies are related to gestational diabetes mellitus in the Korean population. Diabetologia. 2009;52(2):253-61.

80. Kwak SH, Kim SH, Cho YM, Go MJ, Cho YS, Choi SH, et al. A genome-wide association study of gestational diabetes mellitus in Korean women. Diabetes. 2012;61(2):531-41.

81. Wang Y, Nie M, Li W, Ping F, Hu Y, Ma L, et al. Association of six single nucleotide polymorphisms with gestational diabetes mellitus in a Chinese population. PloS one. 2011;6(11):e26953.

82. Lauenborg J, Grarup N, Damm P, Borch-Johnsen K, Jørgensen T, Pedersen O, et al. Common type 2 diabetes risk gene variants associate with gestational diabetes. The Journal of clinical endocrinology and metabolism. 2009;94(1):145-50.

83. Okada Y, Kubo M, Ohmiya H, Takahashi A, Kumasaka N, Hosono N, et al. Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. Nature genetics. 2012;44(3):302-6.

84. Liang J, Pei Y, Liu X, Qiu Q, Sun Y, Zhu Y, et al. The CDKAL1 gene is associated with impaired insulin secretion and glucose-related traits: the Cardiometabolic Risk in Chinese (CRC) study. Clinical endocrinology. 2015;83(5):651-5.

85. Stancáková A, Pihlajamäki J, Kuusisto J, Stefan N, Fritsche A, Häring H, et al. Single-nucleotide polymorphism rs7754840 of CDKAL1 is associated with impaired insulin secretion in nondiabetic offspring of type 2 diabetic subjects and in a large sample of men with normal glucose tolerance. The Journal of clinical endocrinology and metabolism. 2008;93(5):1924-30.

86. Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nature genetics. 2007;39(6):724-6.

87. Shen F, Huang W, Huang JT, Xiong J, Yang Y, Wu K, et al. Decreased N(6)-methyladenosine in peripheral blood RNA from diabetic patients is associated with FTO expression rather than ALKBH5. The Journal of clinical endocrinology and metabolism. 2015;100(1):E148-54.

88. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316(5826):889-94.

89. Merkestein M, Laber S, McMurray F, Andrew D, Sachse G, Sanderson J, et al. FTO influences adipogenesis by regulating mitotic clonal expansion. Nature communications. 2015;6:6792.

90. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gómez-Marín C, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature. 2014;507(7492):371-5.

91. Bravard A, Vial G, Chauvin MA, Rouillé Y, Bailleul B, Vidal H, et al. FTO contributes to hepatic metabolism regulation through regulation of leptin action and STAT3 signalling in liver. Cell communication and signaling : CCS. 2014;12:4.

92. Guo F, Zhang Y, Zhang C, Wang S, Ni Y, Zhao R. Fatmass and obesity associated (FTO) gene regulates gluconeogenesis in chicken embryo fibroblast cells. Comparative biochemistry and physiology Part A, Molecular & integrative physiology. 2015;179:149-56.

93. Mizuno TM, Lew PS, Luo Y, Leckstrom A. Negative regulation of hepatic fat mass and obesity associated (Fto) gene expression by insulin. Life sciences. 2017;170:50-5.

94. Chang Y-C, Liu P-H, Lee W-J, Chang T-J, Jiang Y-D, Li H-Y, et al. Common variation in the fat mass and obesity-associated (FTO) gene confers risk of obesity and modulates BMI in the Chinese population. Diabetes. 2008;57(8):2245-52.

95. Lauenborg J, Grarup N, Damm P, Borch-Johnsen K, Jørgensen T, Pedersen O, et al. Common type 2 diabetes risk gene variants associate with gestational diabetes. 2009;94(1):145-50.

96. Cho YM, Kim TH, Lim S, Choi SH, Shin HD, Lee HK, et al. Type 2 diabetes-associated genetic variants discovered in the recent genome-wide association studies are related to gestational diabetes mellitus in the Korean population. Diabetologia. 2009;52(2):253-61.

97. Franzago M, Fraticelli F, Nicolucci A, Celentano C, Liberati M, Stuppia L, et al. Molecular Analysis of a Genetic Variants Panel Related to Nutrients and Metabolism: Association with Susceptibility to Gestational Diabetes and Cardiometabolic Risk in Affected Women. Journal of Diabetes Research. 2017;2017:4612623.

98. Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, Hewitson KS, et al. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science. 2007;318(5855):1469-72.

99. Gao X, Shin YH, Li M, Wang F, Tong Q, Zhang P. The fat mass and obesity associated gene FTO functions in the brain to regulate postnatal growth in mice. PloS one. 2010;5(11):e14005.

100. Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. The New England journal of medicine. 2015;373(10):895-907.
101. Almawi WY, Nemr R, Keleshian SH, Echtay A, Saldanha FL, AlDoseri FA, et al. A replication study of 19 GWAS-validated type 2 diabetes at-risk variants in the Lebanese population. Diabetes research and clinical practice. 2013;102(2):117-22.
102. Xi B, Takeuchi F, Meirhaeghe A, Kato N, Chambers JC, Morris AP, et al. Associations of genetic variants in/near body mass index-associated genes with type 2 diabetes: a systematic meta-analysis. Clinical endocrinology. 2014;81(5):702-10.
103. Hertel JK, Johansson S, Sonestedt E, Jonsson A, Lie RT, Platou CG, et al. FTO, type 2 diabetes, and weight gain throughout adult life: a meta-analysis of 41,504 subjects from the Scandinavian HUNT, MDC, and MPP studies. Diabetes. 2011;60(5):1637-44.
104. Pagán A, Sabater-Molina M, Olza J, Prieto-Sánchez MT, Blanco-Carnero JE, Parrilla JJ, et al. A gene variant in the transcription factor 7-like 2 (TCF7L2) is associated with an increased risk of gestational diabetes mellitus. European journal of obstetrics, gynecology, and reproductive biology. 2014;180:77-82.
105. Ng MC, Tam CH, So WY, Ho JS, Chan AW, Lee HM, et al. Implication of genetic variants near NEGR1, SEC16B, TMEM18, ETV5/DGKG, GNPDA2, LIN7C/BDNF, MTCH2, BCDIN3D/FAIM2, SH2B1, FTO, MC4R, and KCTD15 with obesity and type 2 diabetes in 7705 Chinese. The Journal of clinical endocrinology and metabolism. 2010;95(5):2418-25.
106. Hu C, Zhang R, Wang C, Wang J, Ma X, Lu J, et al. PPARG, KCNJ11, CDKAL1, CDKN2A-CDKN2B, IDE-KIF11-HHEX, IGF2BP2 and SLC30A8 are associated with type 2 diabetes in a Chinese population. PloS one. 2009;4(10):e7643.
107. Lewis JP, Palmer ND, Hicks PJ, Sale MM, Langefeld CD, Freedman BI, et al. Association analysis in african americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. Diabetes. 2008;57(8):2220-5.
108. Bressler J, Kao WH, Pankow JS, Boerwinkle E. Risk of type 2 diabetes and obesity is differentially associated with variation in FTO in whites and African-Americans in the ARIC study. PloS one. 2010;5(5):e10521.

109. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nature genetics. 2009;41(1):25-34.
110. Xi B, Cheng H, Shen Y, Chandak GR, Zhao X, Hou D, et al. Study of 11 BMI-associated loci identified in GWAS for associations with central obesity in the Chinese children. PloS one. 2013;8(2):e56472.
111. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS genetics. 2007;3(7):e115.
112. Meyre D, Delplanque J, Chèvre JC, Lecoeur C, Lobbens S, Gallina S, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nature genetics. 2009;41(2):157-9.
113. Fredriksson R, Hägglund M, Olszewski PK, Stephansson O, Jacobsson JA, Olszewska AM, et al. The obesity gene, FTO, is of ancient origin, up-regulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain. Endocrinology. 2008;149(5):2062-71.
114. Kelly SJ, Ismail M. Stress and type 2 diabetes: a review of how stress contributes to the development of type 2 diabetes. Annual review of public health. 2015;36:441-62.
115. Lindquist S. The heat-shock response. Annual review of biochemistry. 1986;55:1151-91.
116. Gehrmann M, Marienhagen J, Eichholtz-Wirth H, Fritz E, Ellwart J, Jäättelä M, et al. Dual function of membrane-bound heat shock protein 70 (Hsp70), Bag-4, and Hsp40: protection against radiation-induced effects and target structure for natural killer cells. Cell death and differentiation. 2005;12(1):38-51.
117. Macario AJ, Conway de Macario E. Stress and molecular chaperones in disease. International journal of clinical & laboratory research. 2000;30(2):49-66.
118. Chouchane L, Danguir J, Beji C, Bouassida K, Camoin L, Sfar H, et al. Genetic variation in the stress protein hsp70-2 gene is highly associated with obesity. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity. 2001;25(4):462-6.

119. Ghanayem N, Khamis S, El-Shafie M, El-Sattar El-Ghobashi Y, El-Shazly RM, El-Din Mahmoud Habib M. The potential association of heat shock protein 70-2 + 1267 A/G gene polymorphism with nephropathy in type II diabetic patients. 2014;27(3):589-93.

120. He L, Deng T, Luo H-s. Heat Shock Protein 70 Gene Polymorphisms and Cancer Risk: A Meta-Analysis. The Scientific World Journal. 2014;2014:540309.

121. Pociot F, Rønningen KS, Nerup J. Polymorphic analysis of the human MHC-linked heat shock protein 70 (HSP70-2) and HSP70-Hom genes in insulin-dependent diabetes mellitus (IDDM). Scandinavian journal of immunology. 1993;38(5):491-5.

122. Umapathy D, Krishnamoorthy E, Muthukumaran PJJoDFC. Is HSP70-hom (C2437T) Single Nucleotide Polymorphism (SNP) asso-ciated with Diabetic Foot Ulcer (DFU) among South Indian population? 2012;4:57-62.

123. Singh R, Kølvraa S, Bross P, Jensen UB, Gregersen N, Tan Q, et al. Reduced heat shock response in human mononuclear cells during aging and its association with polymorphisms in HSP70 genes. Cell stress & chaperones. 2006;11(3):208-15.

124. Favatier F, Bornman L, Hightower LE, Günther E, Polla BS. Variation in hsp gene expression and Hsp polymorphism: do they contribute to differential disease susceptibility and stress tolerance? Cell stress & chaperones. 1997;2(3):141-55.

125. Blaschke F, Takata Y, Caglayan E, Law RE, Hsueh WA. Obesity, peroxisome proliferator-activated receptor, and atherosclerosis in type 2 diabetes. Arteriosclerosis, thrombosis, and vascular biology. 2006;26(1):28-40.

126. Forman BM, Chen J, Evans RM. Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors alpha and delta. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(9):4312-7.

127. Lehrke M, Lazar MA. The many faces of PPARgamma. Cell. 2005;123(6):993-9.

128. Kim HI, Ahn YH. Role of peroxisome proliferator-activated receptor-gamma in the glucose-sensing apparatus of liver and beta-cells. Diabetes. 2004;53 Suppl 1:S60-5.

129. Miles PD, Barak Y, He W, Evans RM, Olefsky JM. Improved insulin-sensitivity in mice heterozygous for PPAR-gamma deficiency. The Journal of clinical investigation. 2000;105(3):287-92.
130. Deeb SS, Fajas L, Nemoto M, Pihlajamäki J, Mykkänen L, Kuusisto J, et al. A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. Nature genetics. 1998;20(3):284-7.
131. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nature genetics. 2000;26(1):76-80.
132. Mori H, Ikegami H, Kawaguchi Y, Seino S, Yokoi N, Takeda J, et al. The Pro12 -->Ala substitution in PPAR-gamma is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. Diabetes. 2001;50(4):891-4.
133. Chauhan G, Spurgeon CJ, Tabassum R, Bhaskar S, Kulkarni SR, Mahajan A, et al. Impact of common variants of PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, and CDKAL1 on the risk of type 2 diabetes in 5,164 Indians. Diabetes. 2010;59(8):2068-74.
134. Heude B, Pelloux V, Forhan A, Bedel JF, Lacorte JM, Clément K, et al. Association of the Pro12Ala and C1431T variants of PPARgamma and their haplotypes with susceptibility to gestational diabetes. The Journal of clinical endocrinology and metabolism. 2011;96(10):E1656-60.
135. Shaat N, Lernmark A, Karlsson E, Ivarsson S, Parikh H, Berntorp K, et al. A variant in the transcription factor 7-like 2 (TCF7L2) gene is associated with an increased risk of gestational diabetes mellitus. Diabetologia. 2007;50(5):972-9.
136. Tankó LB, Siddiq A, Lecoeur C, Larsen PJ, Christiansen C, Walley A, et al. ACDC/adiponectin and PPAR-gamma gene polymorphisms: implications for features of obesity. Obesity research. 2005;13(12):2113-21.
137. Yi F, Brubaker PL, Jin T. TCF-4 mediates cell type-specific regulation of proglucagon gene expression by beta-catenin and glycogen synthase kinase-3beta. The Journal of biological chemistry. 2005;280(2):1457-64.

138. da Silva Xavier G, Loder MK, McDonald A, Tarasov AI, Carzaniga R, Kronenberger K, et al. TCF7L2 regulates late events in insulin secretion from pancreatic islet beta-cells. Diabetes. 2009;58(4):894-905.

139. Shu L, Matveyenko AV, Kerr-Conte J, Cho JH, McIntosh CH, Maedler K. Decreased TCF7L2 protein levels in type 2 diabetes mellitus correlate with downregulation of GIP- and GLP-1 receptors and impaired beta-cell function. Human molecular genetics. 2009;18(13):2388-99.

140. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007;445(7130):881-5.

141. Ding M, Chavarro J, Olsen S, Lin Y, Ley SH, Bao W, et al. Genetic variants of gestational diabetes mellitus: a study of 112 SNPs among 8722 women in two independent populations. Diabetologia. 2018;61(8):1758-68.

142. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. Journal of molecular medicine (Berlin, Germany). 2007;85(7):777-82.

143. Gloyn AL, Weedon MN, Owen KR, Turner MJ, Knight BA, Hitman G, et al. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. Diabetes. 2003;52(2):568-72.

144. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nature genetics. 2006;38(3):320-3.

145. Ezzidi I, Mtiraoui N, Cauchi S, Vaillant E, Dechaume A, Chaieb M, et al. Contribution of type 2 diabetes associated loci in the Arabic population from Tunisia: a case-control study. BMC Med Genet. 2009;10:33.

146. van Hoek M, Dehghan A, Witteman JC, van Duijn CM, Uitterlinden AG, Oostra BA, et al. Predicting type 2 diabetes based on polymorphisms from genome-wide association studies: a population-based study. Diabetes. 2008;57(11):3122-8.

147. Wen J, Rönn T, Olsson A, Yang Z, Lu B, Du Y, et al. Investigation of type 2 diabetes risk alleles support CDKN2A/B, CDKAL1, and TCF7L2 as susceptibility genes in a Han Chinese cohort. PloS one. 2010;5(2):e9153.

148. Včelák J, Vejražková D, Vaňková M, Lukášová P, Bradnová O, Hálková T, et al. T2D risk haplotypes of the TCF7L2 gene in the Czech population sample: the association with free fatty acids composition. Physiological research. 2012;61(3):229-40.

149. Sanghera DK, Ortega L, Han S, Singh J, Ralhan SK, Wander GS, et al. Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. BMC Med Genet. 2008;9:59.

150. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nature genetics. 2010;42(7):579-89.

151. Ng MC, Shriner D, Chen BH, Li J, Chen WM, Guo X, et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. PLoS genetics. 2014;10(8):e1004517.

152. Klein K, Haslinger P, Bancher-Todesca D, Leipold H, Knöfler M, Handisurya A, et al. Transcription factor 7-like 2 gene polymorphisms and gestational diabetes mellitus. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet. 2012;25(9):1783-6.

153. Panfili E, Mondanelli G, Orabona C, Belladonna ML, Gargaro M, Fallarino F, et al. Novel mutations in the WFS1 gene are associated with Wolfram syndrome and systemic inflammation. Human molecular genetics. 2021;30(3-4):265-76.

154. Ishihara H, Takeda S, Tamura A, Takahashi R, Yamaguchi S, Takei D, et al. Disruption of the WFS1 gene in mice causes progressive beta-cell loss and impaired stimulus-secretion coupling in insulin secretion. Human molecular genetics. 2004;13(11):1159-70.

155. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, et al. Common variants in WFS1 confer risk of type 2 diabetes. Nature genetics. 2007;39(8):951-3.

156. Franks PW, Rolandsson O, Debenham SL, Fawcett KA, Payne F, Dina C, et al. Replication of the association between variants in WFS1 and risk of type 2 diabetes in European populations. Diabetologia. 2008;51(3):458-63.
157. Cheurfa N, Brenner GM, Reis AF, Dubois-Laforgue D, Roussel R, Tichet J, et al. Decreased insulin secretion and increased risk of type 2 diabetes associated with allelic variations of the WFS1 gene: the Data from Epidemiological Study on the Insulin Resistance Syndrome (DESIR) prospective study. Diabetologia. 2011;54(3):554-62.
158. Florez JC, Jablonski KA, McAteer J, Sandhu MS, Wareham NJ, Barroso I, et al. Testing of diabetes-associated WFS1 polymorphisms in the Diabetes Prevention Program. Diabetologia. 2008;51(3):451-7.
159. Sparsø T, Andersen G, Albrechtsen A, Jørgensen T, Borch-Johnsen K, Sandbaek A, et al. Impact of polymorphisms in WFS1 on prediabetic phenotypes in a population-based sample of middle-aged people with normal and abnormal glucose regulation. Diabetologia. 2008;51(9):1646-52.
160. Fonseca SG, Fukuma M, Lipson KL, Nguyen LX, Allen JR, Oka Y, et al. WFS1 is a novel component of the unfolded protein response and maintains homeostasis of the endoplasmic reticulum in pancreatic beta-cells. The Journal of biological chemistry. 2005;280(47):39609-15.
161. Holst JJ, Vilsbøll T, Deacon CF. The incretin system and its role in type 2 diabetes mellitus. Molecular and cellular endocrinology. 2009;297(1-2):127-36.
162. Takeda K, Inoue H, Tanizawa Y, Matsuzaki Y, Oba J, Watanabe Y, et al. WFS1 (Wolfram syndrome 1) gene product: predominant subcellular localization to endoplasmic reticulum in cultured cells and neuronal expression in rat brain. Human molecular genetics. 2001;10(5):477-84.
163. Barrett TG, Bundey SE. Wolfram (DIDMOAD) syndrome. Journal of medical genetics. 1997;34(10):838-41.
164. Riggs AC, Bernal-Mizrachi E, Ohsugi M, Wasson J, Fatrai S, Welling C, et al. Mice conditionally lacking the Wolfram gene in pancreatic islet beta cells exhibit diabetes as a result of enhanced endoplasmic reticulum stress and apoptosis. Diabetologia. 2005;48(11):2313-21.
165. Yamada T, Ishihara H, Tamura A, Takahashi R, Yamaguchi S, Takei D, et al. WFS1-deficiency increases endoplasmic reticulum stress, impairs cell cycle

progression and triggers the apoptotic pathway specifically in pancreatic beta-cells. Human molecular genetics. 2006;15(10):1600-9.

166. Dupépé EB, Kicielinski KP, Gordon AS, Walters BC. What is a Case-Control Study? Neurosurgery. 2019;84(4):819-26.

167. Kicielinski KP, Dupépé EB, Gordon AS, Mayo NE, Walters BC. What Isn't a Case-Control Study? Neurosurgery. 2019;84(5):993-9.

168. Das S, Behera MK, Misra S, Baliarsihna AK. Beta-cell function and insulin resistance in pregnancy and their relation to fetal development. Metabolic syndrome and related disorders. 2010;8(1):25-32.

169. Saucedo R, Zarate A, Basurto L, Hernandez M, Puello E, Galvan R, et al. Relationship between circulating adipokines and insulin resistance during pregnancy and postpartum in women with gestational diabetes. Archives of medical research. 2011;42(4):318-23.

170. Uebel K, Pusch K, Gedrich K, Schneider K-TM, Hauner H, Bader BL. Effect of maternal obesity with and without gestational diabetes on offspring subcutaneous and preperitoneal adipose tissue development from birth up to year-1. BMC Pregnancy and Childbirth. 2014;14(1):138.

171. Mørkrid K, Jenum AK, Berntsen S, Sletner L, Richardsen KR, Vangen S, et al. Objectively recorded physical activity and the association with gestational diabetes. Scandinavian journal of medicine & science in sports. 2014;24(5):e389-97.

172. Alam M, Jahan S, Hasan M, Sultana N, Mohona Z, Rakibul-Hasan M, et al. Insulin Secretory Defect May Be the Major Determinant of GDM in Lean Mothers. Journal of Clinical & Translational Endocrinology. 2020;20:100226.

173. Mahjabeen S, Rahman, A. , Hossain, M. , Mustari, M. , Abedin, M. , Mahmood, T. , Chowdhury, T. , Islam, T. and Hossain, R. Insulin Sensitivity and Insulin Secretion Estimated by Homeostatic Model Assessment (HOMA) in Gestational Diabetes Mellitus. . Journal of Biosciences and Medicines. 2020;8:44-54.

174. Crow JF. Eighty years ago: the beginnings of population genetics. Genetics. 1988;119(3):473-6.

175. Robertson A, Hill WG. Deviations from Hardy-Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. Genetics. 1984;107(4):703-18.

176. Shete JWaS. Testing Departure from Hardy–Weinberg Proportions. In: Sun RCEMS, editor. Statistical Human Genetics: Methods and Protocols. 850: Springer Nature 2012.

177. Gordon D, Ott J. Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. Biocomputing 2001: World Scientific; 2000. p. 18-29.

178. Tiret L, Cambien FJC. Departure from Hardy-Weinberg equilibrium should be systematically tested in studies of association between genetic markers and disease. 1995;92(11):3364-5.

179. Xu J, Turner A, Little J, Bleecker ER, Meyers DAJHg. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? 2002;111(6):573.

180. Hosking L, Lumsden S, Lewis K, Yeo A, McCarthy L, Bansal A, et al. Detection of genotyping errors by Hardy–Weinberg equilibrium testing. 2004;12(5):395-9.

181. Attia J, Thakkinstian A, McElduff P, Milne E, Dawson S, Scott RJ, et al. Detecting genotyping error using measures of degree of Hardy-Weinberg disequilibrium. 2010;9(1).

182. Chen B, Cole JW, Grond-Ginsbach C. Departure from Hardy Weinberg Equilibrium and Genotyping Error. 2017;8.

183. Verani JR, Baqui AH, Broome CV, Cherian T, Cohen C, Farrar JL, et al. Case-control vaccine effectiveness studies: Preparation, design, and enrollment of cases and controls. Vaccine. 2017;35(25):3295-302.

184. VanderWeele TJ. Principles of confounder selection. European journal of epidemiology. 2019;34(3):211-9.

3. Methods and Materials

3.1 Study Design

The basic steps of this study are shown in **figure 3.1**. It is a case-control study and started with selecting genes related to T2DM and target SNP selection. This was followed by sample collection from pregnant women for genotyping and association analyses of selected SNPs with GDM.

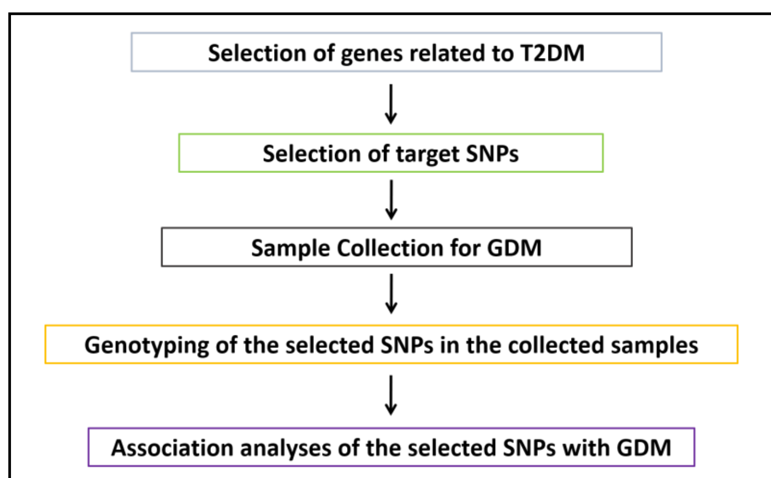
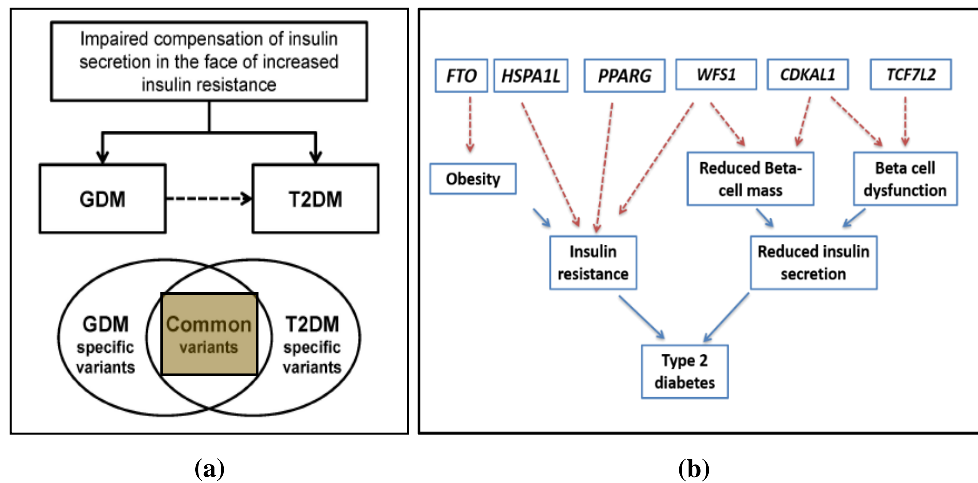


Figure 3.1: Steps of the overall study

3.1.1 Selection of target genes:

T2DM and GDM share common pathophysiological backgrounds like insulin resistance and impaired compensatory insulin secretion; there must be some genetic similarities between these two diseases [**Figure 3.2(a)**] (1). These common variants are the cornerstones of this study. As insulin resistance or reduced insulin secretion or combination of both results T2DM, associations of the variants of these genes with the pathophysiology of this disease have been emphasized during gene selection (1-7).



(a) (b)

Figure 3.2: (a) GDM and T2DM have common pathophysiology adapted from (1) (b) Schematic diagram of selected genes variants of which related to the pathophysiology of T2DM

Six genes, variants of which related to one or both the pathophysiology, have been selected and shown in figure 3.2 (b)(2, 6-9). These genes are listed in table 3.1.

Table 3.1: List of genes used for SNP selection in this study

Sl. No	Gene Name	Chromosome
1	CDK5 Regulatory Subunit Associated Protein 1-Like1(***CDKAL1)***	6
2	Fat mass and obesity associated gene (***FTO)***	16
3	Heat Shock Protein Family A (Hsp70) Member 1 Like ***(HSPA1L)***	6
4	Peroxisome proliferator-activated receptor gamma (***PPARG***)	3
5	Transcription factor 7 like 2 (***TCF7L2***)	10
6	Wolfram syndrome 1 (***WFS1***)	4

3.1.2 Selection of target SNPs:

Ten SNPs (**Table 3.2**) of the six genes have been selected on the basis of the literature review (**Section 2.7-2.12**). Eight of them are intron variants, and two are coding sequence variants (*HSPA1L* gene variant rs2227956 and *PPARG* rs1801282). All are reported for their association with T2DM and, in some cases, with GDM in different populations (3-5).

Table 3.2: List of selected SNPs and reported T2DM risk by their reference and altered alleles

Gene	SNP Id	Normal allele	Hetero allele	Increase T2DM risk	Minor allele	Increase T2DM risk	Reference
CDKAL1	rs7756992	A/A	A/G	1.3×	G/G	1.3×	(10, 11)
	rs7754840	G/G	C/G	1.3×	C/C	1.3×	(10, 11)
TCF7L2	rs7903146	C/C	C/T	1.4×	T/T	2 ×	(11-13)
	rs12255372	G/G	G/T	1.3×	T/T	1.5×	(11-13)
	rs10885406	A/A	A/G	1.2×	G/G	1.8×	(14, 15)
PPARG	rs1801282	C/C	C/G	1.2×	G/G	1.8×	(8)
	rs3856806	C/C	C/T	1.1×	T/T	1.5×	(16, 17)
WFS1	rs10010131	A/A	A/G	1.2×	G/G	1.4×	(18)
FTO	rs8050136	C/C	A/C	1.2×	A/A	1.4×	(19, 20)
HSPA1L	rs2227956	T/T	C/T	1.8×	C/C	2.1×	(9)

Source: https://www.snpedia.com/index.php/SNPedia

3.1.3 Sample collection

Participants of this study were recruited from the Department of Gynecology and Obstetrics, Bangabandhu Sheikh Mujib Medical University (BSMMU). After assessing their eligibility, they were requested to visit the Department of Endocrinology and Metabolism, BSMMU, to screen and diagnose GDM. Participants were divided into two groups after screening and diagnosis. Blood samples were collected from both groups. Participants diagnosed with GDM were advised to follow up screening after delivery (**Figure 3.3**). Participants who had normal glucose levels after delivery confirmed the occurrence of GDM during pregnancy.

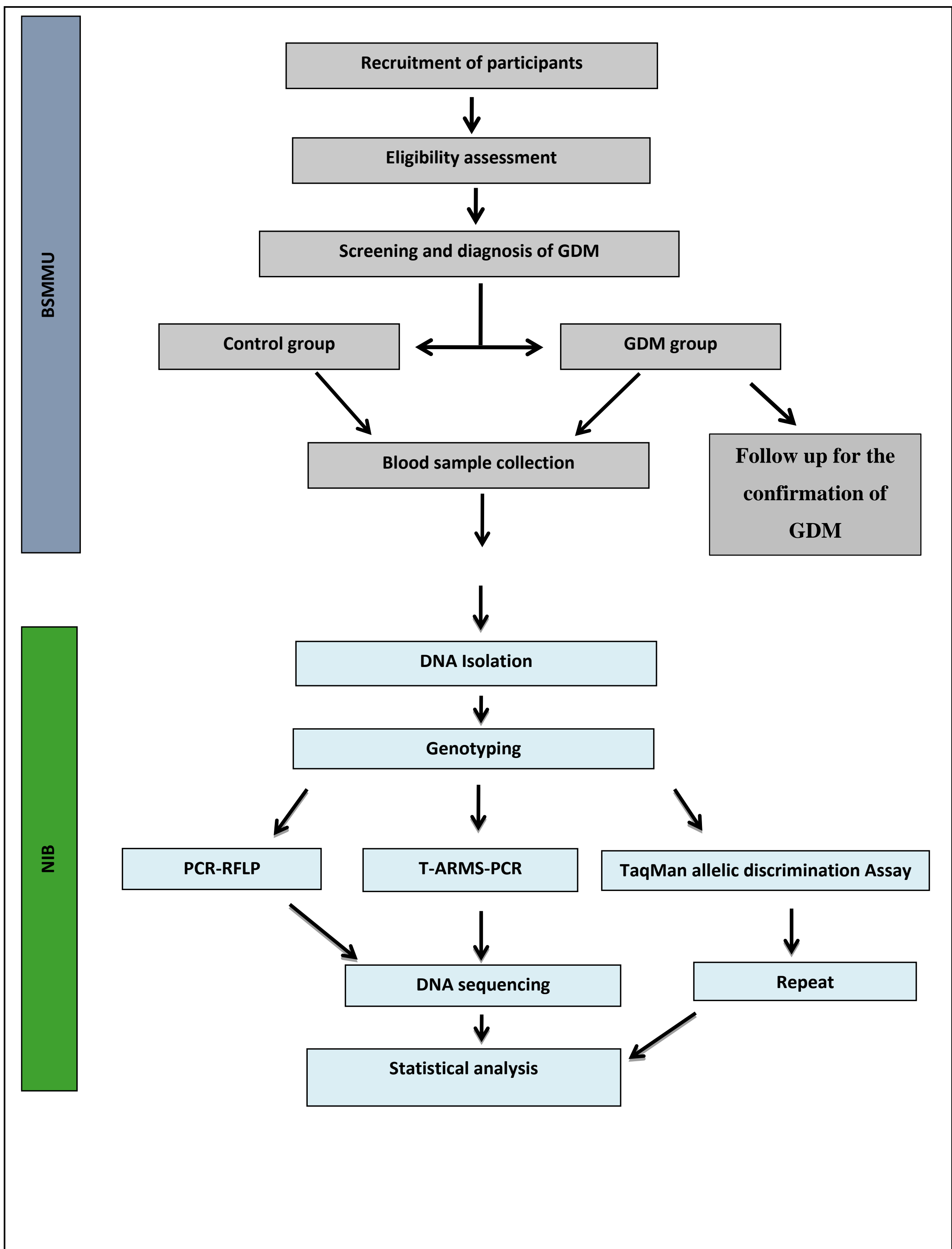


Figure 3.3: Flow chart of the overall methods followed in this study.

3.1.4 Genotyping

Three methods, Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP), Tetra-primer Amplification Refractory Mutation System Polymerase Chain Reaction (T-ARMS-PCR), and TaqMan allelic discrimination assay, were used for genotyping. Results obtained by RFLP and T-ARMS PCR were confirmed by DNA sequencing of blindly selected samples, and those by TaqMan allelic discrimination assay were confirmed by repeated TaqMan assay of a portion of the total samples (**Figure 3.3**).

3.1.5 Association analyses

Finally, association of selected SNPs with GDM was assessed by statistical analyses.

3.2 Screening and enrollment of the participants

3.2.1 Recruitment of Study subjects

Participants of this study were recruited irrespective of the trimester. If the mother's glycemic status was found to be normal before the 24th week of pregnancy, she was instructed to repeat the OGTT between the 24th and 28th weeks of pregnancy, and the glycemic status was reassessed (**Figure 3.4**). Otherwise, they were not included in this research.

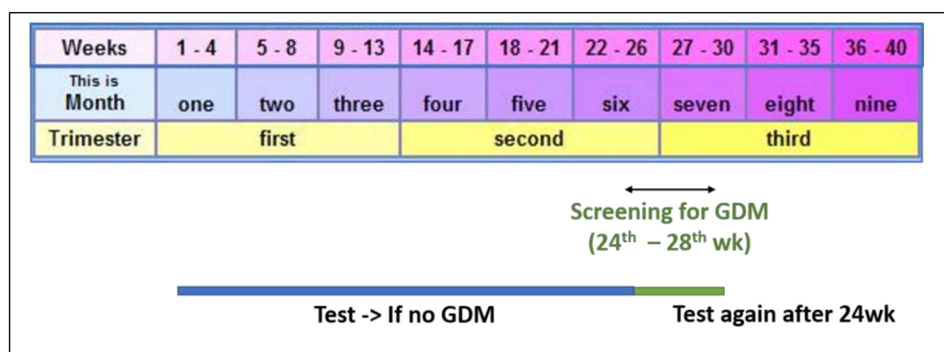


Figure 3.4: Screening for selecting study participants.

3.2.2 Eligibility Assessment

3.2.2.1 Inclusion criteria

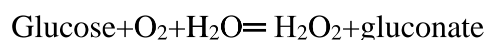
- Women with pregnancy of any duration
- Pregnant mothers giving consent for study

3.2.2.2 Exclusion criteria

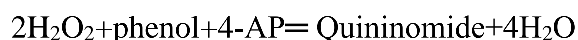
- Diabetes in pregnancy (Pregestational hyperglycemia)
- Steroid treatment recipient
- Thyroid dysfunction
- Acute critical illness
- Chronic liver and kidney disease
- Heart disease
- Known diabetes mellitus
- Pregnant mothers not giving consent to take part in the study.

3.3 Diagnosis of GDM by Oral Glucose Tolerance Test (OGTT)

Participants underwent 75g OGTT on the appointed day after 8-10 hours of overnight fast, and glycemic status was determined using WHO 2013 criteria for GDM. Blood samples were collected by venipuncture three times during fast and after 1 and 2 hours of 75g glucose loads. Sodium fluoride containing gray-top tubes were used for blood sample collection and plasma were separated within 45 minutes. Plasma glucose was assayed by the glucose oxidase method using Dimension EXL 200 Integrated Chemistry System in an automated analyzer (Siemens, Germany) at the Department of Biochemistry, BSMMU. Here glucose is oxidized by glucose oxidase to gluconate and hydroxy peroxide according to the following equation:



$$Glucose+O_2+H_2O = H_2O_2+gluconate$$



$$2H_2O_2+phenol+4\text{-}AP = Quininomide+4H_2O$$

Reagent-1 (TRIS buffer PH 7.5+phenol) and reagent-2 (glucose oxidase, peroxidase, 4-aminophenazene) were mixed to form a solution that remains stable at 2-8°C for one month. Plasma was added and incubated at 37°C for 10 minutes to yield a colorimetric

reaction (at 550 nm wavelength). The result was deducted from computerized calculation utilizing standard curve derived from known concentrations used by the system.

WHO (2013) recommend a 75gm OGTT and the diagnosis of GDM is made irrespective of weeks of gestation:

Table 3.3: WHO 2013 Criteria

Time points	GDM	DM in pregnancy (DIP)
Fasting PG	5.1-6.9 mmol/L	≥ 7.0 mmol/L
01-hour PG	≥ 10.0 mmol/L	-
02-hour PG	8.5-11.0 mmol/L	≥ 11.1 mmol/L

PG: plasma glucose

If at least one value of plasma glucose concentration is equal to or exceeds thresholds the individual was diagnosed with GDM. Pregnant women with all 3 plasma glucose values during 75 gm OGTT below the WHO (2013) recommended cut off for diagnosis of GDM were considered as normoglycemic controls (**Table 3.3**).

3.4 Insulin Indices

3.4.1 Fasting serum insulin: an index of insulin resistance

3.4.2 Homeostasis Model Assessment (HOMA)

- HOMA of insulin resistance (HOMA-IR) index

HOMA-IR was calculated using the following formula:

$$\text{HOMA-IR} = \frac{\text{Fasting insulin } (\mu\text{IU/ml}) \times \text{Fasting glucose (mmol/L)}}{22.5}$$

- HOMA of β-cell function (HOMA-B) index

$$\text{HOMA-B} = \frac{20 \times \text{Fasting insulin } (\mu\text{IU/ml})}{\text{Fasting glucose (mmol/L)} - 3.5}$$

- HOMA of insulin sensitivity (HOMA-%S) index

$$\text{HOMA-\%S} = \frac{1}{\text{HOMA-IR}} \times 100$$

In this study, HOMA-IR values below 2.89 was considered as normal, whereas values equal to or above 2.89 as IR (21).

The Access Immunoassay System (REF- 33410), Beckman Coulter, Inc., USA, was used to quantitatively measure serum insulin levels utilizing a chemiluminescent immunoassay approach.

3.4.3 Principles of the Procedure

The Access Ultrasensitive Insulin assay is a simultaneous one-step immune enzymatic ("sandwich") assay. Samples were added to reaction vessels along with mouse monoclonal anti-insulin alkaline phosphatase conjugate and paramagnetic particles coated with mouse monoclonal anti-insulin antibody. The serum insulin bound to the antibody on the solid phase, while the conjugate reacted with a different antigenic site on the insulin molecule. After incubation in a reaction vessel, materials bound to the solid phase were held in a magnetic field while unbound materials were washed away. Then, the chemiluminescent substrate Lumi-Phos 530 was added to the vessels and light generated by the reaction was measured with a luminometer. The light production was directly proportional to the concentration of insulin in the sample. The amount of analyte in the sample was determined from a stored, multi-point calibration curve.

3.5 Data collection Procedure:

A structured data collection sheet was used for this purpose (**Appendix A1**). Demographic and anthropometric measures as well as other information of all study subjects were recorded in the data collection sheet.

3.5.1 Demographic and anthropometric data collection

On the day of OGTT, anthropometric measurements of participants, including weight, height, and blood pressure, were measured. Demographic data were recorded after asking the participants specific questions. Bodyweight was measured using a calibrated

digital scale for weight. Height was measured using a mounted measuring tape for height measurement (with precision to 0.1 cm). Blood pressure was measured using a calibrated sphygmomanometer (ALPK2 500-V, Japan) in a sitting position after relaxation for at least 15 min.

3.5.2 Body mass index

The BMI was computed by dividing the weight in kilograms by the square of the height in meters. Both GDM patients and pregnant normoglycemic controls had their BMI measured.

3.5.3 Gravida

Number of conceptions by our study population including the current pregnancy. It includes miscarriage, abortion, still birth and events of menstrual regulation.

3.5.4 Occupational status

A study participant who was not directly involved in income generation activities was designated as "Housewife." Women directly involved in income generation activities were referred to as "service holders." "Others" are the participants who were not in either of these two groups.

3.6 Blood Sample Collection

After screening and diagnosis of GDM, study participants were divided into two groups; normoglycemic controls and GDM cases. Blood samples were collected from both groups during OGTT and collected in BD Vacutainer® K2 EDTA (BD Franklin Lakes NJ USA) tubes. Blood samples were transported to NIB for genotyping, maintaining appropriate temperature, and stored at -40°C for future use.

3.7 DNA Extraction

According to the manufacturer's protocol, genomic DNA was extracted from blood samples using the PureLink® Genomic DNA extraction kit (Invitrogen). The protocol used for DNA isolation is described in Appendix A2. The NanoDrop 2000 UV Vis Spectrophotometer was used to determine the purity and concentration of the isolated

DNA. For genotyping, DNA samples with an OD260/OD280 ratio of 1.8 to 2.0 and concentrations of greater than 80 ng/mL were employed.

3.8 Quality assessment and quantification of DNA

The quality of DNA is very important to obtain good results and for long-term storage. Degradation often occurs due to careless handling. It is also important to know the exact concentration of the DNA for subsequent analyses.

NanoDrop spectrophotometer can measure the concentration of nucleic acid (both DNA and RNA), protein samples and others with only one microliter of sample within a few seconds. It also shows the standard curve (**Figure 3.5**) of the sample for quality assurance. The procedure to measure the nucleic acid concentration is as follows:

- Option for nucleic acid concentration measurement was selected. The wavelength was fixed at 260 and 280 nm for nucleic acid analysis.
- The nozzle of the machine was first cleaned with soft tissue after lifting its lid and was initialized with PCR grade water.

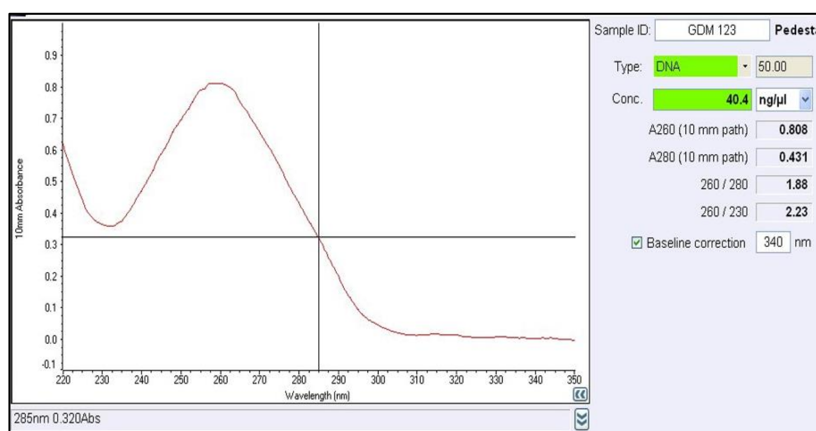


Figure 3.5: Measuring the concentration of DNA using a NanoDrop spectrophotometer

- After initialization, the blank was set with appropriate buffer according to the buffer in which the DNA was dissolved. [Optical density (OD) of buffer was taken as blank].
- Two microliters (µL) of the sample nucleic acid were loaded onto the nozzle, the lid was closed and the OD measured.

3.9 Polymorphism analysis

3.9.1 Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP)

The PCR RFLP is a relatively simple and inexpensive method for SNP genotyping. Five of the target SNPs were located within restriction sites so presence of these SNPs may create or destroy the restriction enzyme binding sites which respectively results digestion or no digestion of the PCR products containing SNP. The resulting restriction fragments are than separated by gel electrophoresis according to their size (**Figure 3.6**).

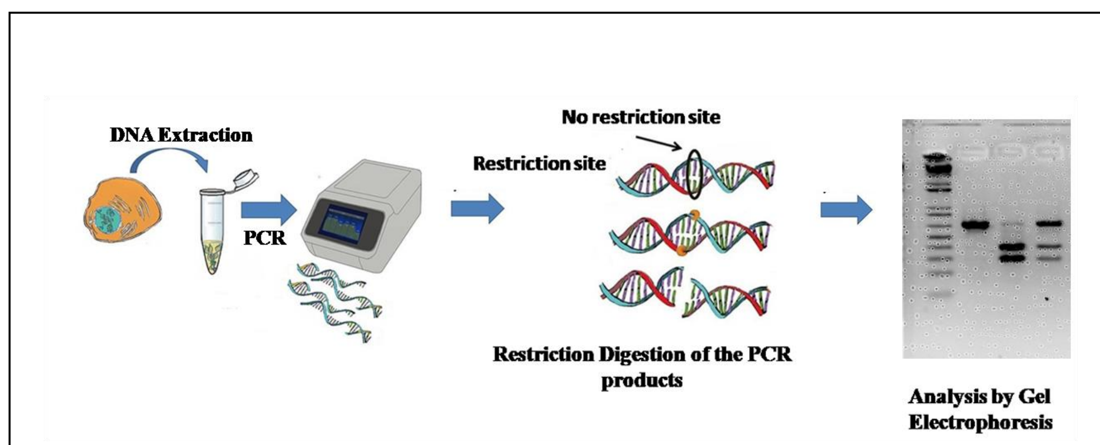


Figure 3.6: SNP genotyping by PCR-Restriction Fragment Length Polymorphism (PCR-RFLP)

3.9.1.1 PCR amplification of regions spanning target SNPs

3.9.1.1.1 Primer Design

Both RFLP method and DNA sequencing required amplification of sequences that spanned the selected SNPs. To amplify the genomic regions surrounds the candidate SNPs, primers were designed from the flanking regions of the SNPs using the sequence information from human genome *GRCh38* assembly (www.ncbi.nlm.nih.gov) (Table 3.4). Primer pairs were designed by using Primer3Plus tool (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and checked by OligoAnalyzer™ Tool (www.idtdna.com/pages/tools/oligoanalyzer). The primers were diluted in TE buffer.

Table 3.4: List of primers used to amplify regions spanning variants genotyped by PCR-RFLP method

Gene	SNP	Primer	Tm (°C)	Optimized Tm (°C)
TCF7L2	rs12255372	F: 5′-CTGGAAACTAAGGCGTGAGG-3′	54	50
		R: 5′-ATGCCACCCAAGGTTTGA-3′	48	
CDKAL1	rs7756992	F: 5′-TTGATTGTAAAGACTGGGTCTCA-3′	52	50
		R: 5′-GAACGAAGGCAAATAAATTCAA-3′	47	
PPARG	rs3856806	F: 5′-TTACATCCTGGCCAGAAAAA-3′	48	50
		R: 5′ TGCTTTTTCACAGTAAATTTCTTAGG-3′	52	
WFS1	rs10010131	F: 5′-ACCTCTGAGAGAGGGGAGGA-3′	56	55
		R: 5′-TAGGGCACGGTCTCTACAGG-3′	56	
HSPA1L	rs2227956	F: 5′-GGACAAGTCTGAGAAGGTACAG-3′	55	61
		R: 5′-GTAACTTAGATTCAGGTCTGG-3′	50	

- Primer sequence to amplify *HSPA1L* gene variant's spanning region was designed and characterized by Moniruzzaman et. al.,2020 (9).

Table 3.5: List of primers used for sequencing

Gene	SNP	Primer	Tm (°C)	Optimized Tm (°C)
TCF7L2	rs10885406	F: 5′-TGTGGCCTATTGCAGTTGAG-3′	52	50
		R: 5′-AATCAGGGGCATGCATTAAA-3′	48	
	rs7903146	F: 5′-TGAAGACATACACAAAAGTTTTATTGG-3′	52	55
		R: 5′-CAGAATGAGACCCTGTCTCTGA-3′	55	
CDKAL1	rs7754840	F: 5′-GTGTTTGGCCTTGAGTTTGG-3′	52	55
		R: 5′-CTGCTCACTGGCATACATCA-3′	52	
PPARG	rs1801282	F: 5′-GCCCCTCACAAGACACTGA-3′	53	55
		R: 5′-CCTGGAAGACAAACTACAAGAGC-3′	55	

3.9.1.1.2 PCR amplification

PCR reactions were carried out with 80 ng of DNA, 1× GoTaq® G2 Hot Start Master Mix (Promega, USA), 10µM of each primer and H_2O. PCR program was carried out as follows: initial denaturation at 95˚C for 5 min, followed by 35 cycles of 45 sec at 95˚C, 30 sec at respective annealing temperature and 45 sec at 72˚C, then a final extension of 7 min at 72˚C.

3.9.1.1.3 Analysis of PCR Products

The fraction of the amplified products was analyzed by agarose gel electrophoresis (**Appendix A3**) and length of the PCR product was determined by using 1kb+ DNA marker (Invitrogen, USA). For confirmation of PCR amplification, a 1% agarose gel

was always used. The preparation of 1% Agarose gel and maintenance of gel electrophoresis reagents were outlined at appendix A3. If the product size was perfectly matched with the assumed one the subsequent restriction digestion was carried out for genotyping.

3.9.1.2 Restriction digestion

After confirmation of accurate amplification of the target region, 7µL of the amplified product was digested with appropriate restriction enzyme (NEB) (**Table 3.6**). The restriction endonucleases for RFLP were chosen using online tool NEB cutter (https://nc2.neb.com/NEBcutter2/). RFLP reaction mixtures (10µL) contained 1.5µL of nuclease free water, 7µL of PCR product, 1µL of 10× reaction buffer, and 0.5µL of endonuclease enzyme. Reactions were carried out at specific incubation temperature and time listed in table 3.6.

Table 3.6: PCR product size and restriction enzymes used for PCR-RFLP method

Gene	SNP	PCR Product Size	Restriction Enzyme used	Incubation		Allele	Product size after digestion
				Temp	Time		
CDKAL1	rs7756992	684 bp	BglII	37 °C	O\N	A	388 bp, 296 bp
						G	684 bp
HSPA1L	rs2227956	862 bp	NcoI	37 °C	2hr	T	616 bp, 246 bp
						C	862 bp
PPARG	rs3856806	580 bp	BsaAI	37 °C	1hr	C	396 bp, 184 bp
						T	580 bp
TCF7L2	rs12255372	376 bp	MluCI	37 °C	3hr	G	143bp, 134bp, 99bp
						T	134 bp, 126 bp, 99 bp, 17bp
WFS1	rs10010131	654 bp	BsmF1	37 °C	2hr	A	552 bp, 102 bp
						G	288 bp, 264 bp,102 bp

Temp, Temperature; bp, Base pair; hr, hour; O\N, overnight.

3.9.1.3 Analyses of Polymorphisms from gel electrophoresis of digestion products

For separating different sizes of digestion products depending on the size of the DNA band, different concentrations of agarose were used (22). The digestion products of the *TCF7L2* and *WFS1* gene polymorphisms rs12255372 and rs10010131 were too close to separate by agarose gel and were separated by Polyacrylamide gel electrophoresis (PAGE).

3.9.1.3.1 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel has much higher resolution than the agarose gel. Depending on the polymorphic nature different concentrations (6-12%) of non-denaturing PAGE were used for easy analyzing or scoring. This was done according to following protocol modified from Santos, F.(23). The composition and maintenance of PAGE is given in Appendix A4.

3.9.1.3.1.1 Preparation and procedure of 12% non-denaturing polyacrylamide gel electrophoresis (PAGE)

Components:

To prepare 175 ml polyacrylamide the required ingredients and their volume are given below-

Ingredients	Volume
40% Acrylamide	52.5 ml
25×TBE	8.4 ml
ddH_2O	up to175 ml

↓

Stir for 10 minutes

↓

10% APS	2.1ml
TEMED	178.5μl

↓

Stir for 1 minute

Procedure

1. Both glass plates were cleaned carefully by 99% ethanol and were assembled, placing both cleaned surfaces inside, with spacers (~1.5 mm thick) and elastic rubber as a sealer surrounding the edges of glass plate.

2. The assembly was levelled and checked for leakage with ddH_2O.

3. The gel was poured in the gel case and the comb was assembled for wells formation. The gel was allowed to solidify for ~20 minutes.

4. The combs were removed and sandwich glass plate/ gel was attached with the electrophoresis apparatus.

5. The PCR product was mixed with appropriate volume of loading buffer

6. 4-6µl of the mix was generally loaded onto 10% non-denaturing polyacrylamide gel very carefully to prevent cross contamination.

7. Electrophoresis was carried out in 1× TBE buffer at 350 Volt up to the time when the bromophenol blue/xylene cyanol travelled a satisfactory distance.

8. The electric current was turn off and both glass plates were disassembled. Elastic rubber and spacers were removed.

3.9.2 TaqMan allelic discrimination assay

SNPs which did not located inside a restriction site were genotyped by TaqMan allelic discrimination assay. The graphical representation of the assay is shown in the **Figure 3.7** adopted from (24). Each assay enables genotyping of individuals for a SNP and consist of two sequence specific primers and two TaqMan minor groove binder (MGB) probes with nonfluorescent quencher. One probe is labelled with VIC dye to detect the allele 1 sequence, the second probe is labelled with FAM dye to detect the allele 2 sequence.

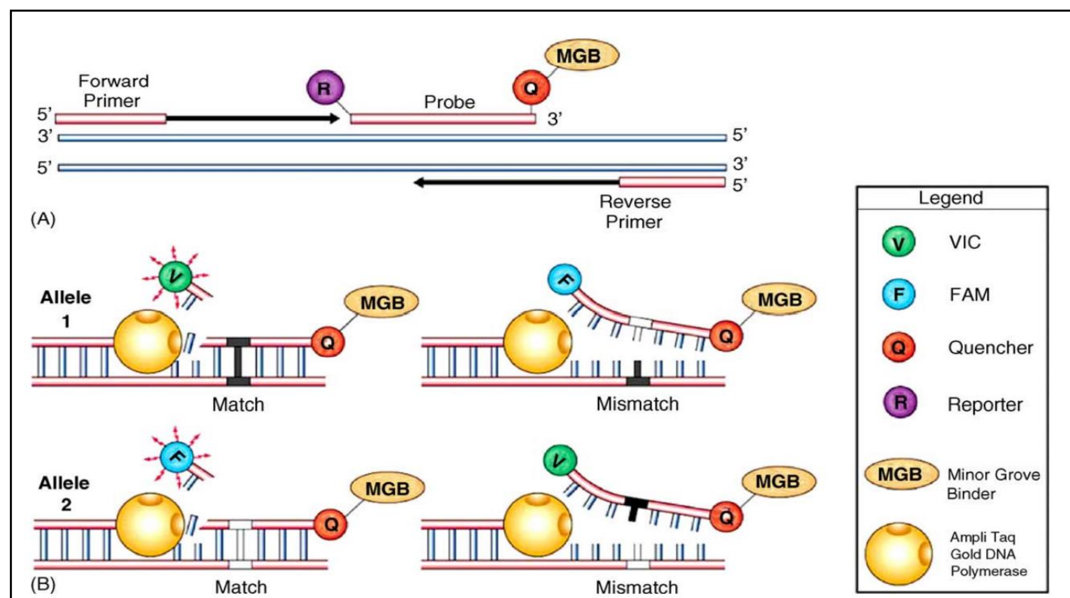


Figure 3.7: Graphical representation of the TaqMan® Genotyping Assay (F.M. De La Vega et al. / Mutation Research 573, 2005)

Procedure of TaqMan assay

1. DNA samples were diluted to a final concentration of 10 ng/μl. Concentration was determined by using Nanodrop (Thermo Fisher).

2. Master mix solution was prepared enough for samples to be run in triplicate, no template control (NTC), three positive controls plus an extra 5% for pipetting error. The reagents and required volume are as follows-

Reagents	Volume
2× TaqMan™ Genotyping Master Mix	12.50 μL
20× Assay mix (working stock) for the respective SNP	1.25 μL
Nuclease-Free H_2O	Up to 25 μL
Total	13.75 μL

3. Mixed well and briefly centrifuged.

4. Required amount of master mix was pipetted into each well of the PCR plate

5. Removed from PCR set-up hood

6. Required volume of DNA was pipetted into all sample wells

7. Three positives (one homozygous wild-type allele carrier, one heterozygous, and one homozygous risk allele carrier) and three negatives (all components excluding DNA) controls were inserted at random in each run as a quality check.

8. Required volumes of positive controls and water for NTC were pipetted into their respective wells.

9. The plate was sealed with adhesive film then centrifuge briefly to bring the reaction mix to the bottom of the well and eliminate air bubbles.

10. Centrifuged briefly and placed on the Quant Studio 5 Real time PCR machine

3.9.3 Tetra-primer Amplification Refractory Mutation System Polymerase Chain Reaction (T-ARMS-PCR)

Tetra primer amplification refractory mutation system PCR uses four primers in a single PCR reaction followed by gel electrophoresis (Figure 3.8). Firstly, two nonallele specific outer primers amplify the region that comprises the SNP. Then two allele

specific inner primers will produce the allele specific fragments. Placing the outer primers at different distances from the SNP, the two allele specific fragments can be distinguished by their different sizes in an agarose gel. The rs8050136 polymorphism of *FTO* gene was genotyped using the T-ARMS-PCR method. Briefly, the region was amplified with the following primers in a single reaction:

Table 3.7: List of primers used in T-ARMS method for the *FTO* gene rs8050136 genotyping

Primer Name	Sequence	Annealing temperature
Outer primer_ F	5′-CTTAAGAGTCCATACCAACCAAGGT-3′	60.96°C
Outer primer_ R	5′-ATAATTGGCTCTCGACATTTACACA-3′	61.07°C
Inner primer_ F	5′-AGTTGCCCACTGTGGCAGTC-3′	63.67°C
Inner primer_ R	5′-GCAAAAACCACAGGCTCAGATACTT-3′	63.93°C

The PCR cycles were as follows: 5 min at 95 °C, followed by 30 cycles of 50 sec at 95 °C, 45 sec at 62 °C, and 45 sec at 72 °C. The final extension was for 7 min at 72 °C. Firstly, two nonallele specific outer primers amplify 337 bp the region that comprises the SNP. Placing the outer primers at different distances from the SNP, the two allele specific fragments can be distinguished by their different sizes in an agarose gel (Figure 3.8).

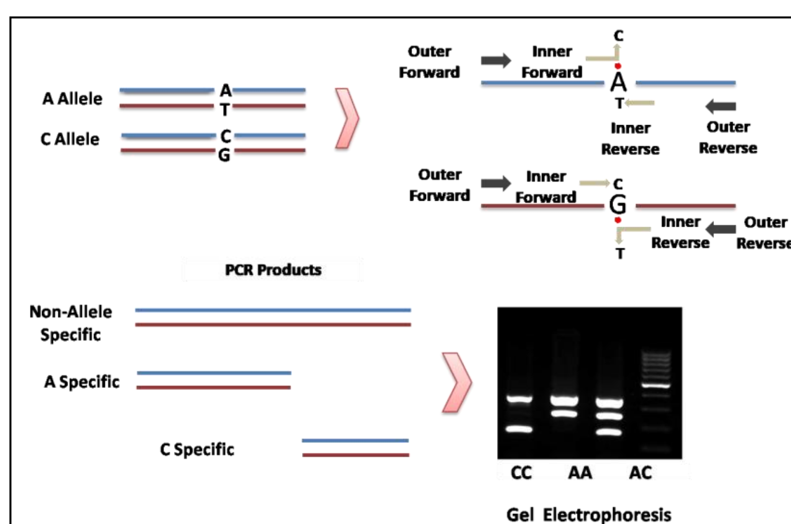


Figure 3.8: Tetra-primer Amplification Refractory Mutation System-Polymerase Chain Reaction (T-ARMS-PCR) for SNP genotyping

Then two allele specific inner primers will produce the allele specific fragments 251 bp and 130 bp. Subsequently, 7µL of the amplified product was subjected to 2% agarose gel electrophoresis and stained with ethidium bromide. The AA genotype was represented by two bands of 337 bp and 251 bp, while the CC genotype was represented by 337 bp and 130 bp products. Three bands were also detected when heterozygous genotype AC was present: 337 bp, 251 bp, and 130 bp.

3.10 DNA sequencing

Genotypes obtained from PCR RFLP and T ARMS PCR were confirmed by DNA sequencing of blindly selected samples. Positive controls for TaqMan genotyping assay were also selected by sequencing.

3.10.1 Template preparation

For sequencing, region spanning the target SNPs were amplified by PCR with specific primer pairs (**Table 3.4 and 3.5**) and checked by agarose gel electrophoresis for appropriate amplification (**section 3.8.1.1.3**). The amplified PCR product was then purified by using the PureLink™ PCR Purification Kit (Invitrogen™) following the manufacturer's instructions.

3.10.2 Cycle sequencing

The purified PCR products were then cycle sequenced using Big Dye terminator V 3.1 reactions mixture (Applied Biosystems, Foster City, CA) and only one primer. The reaction conditions are as follows

Reaction condition

20 µL cycle sequencing reaction mixtures (for >500bp sequencing) were carried out

Sl. No	Reagent	Volume (µL)
1.	5× Sequencing Buffer	4.0
2.	Big Dye® Terminator v3.1	4.0
3.	Template (40 ng/ µL)	1.0
4.	Primer (10 pmol/µL)	0.5
5.	Deionized H_2O	10.5
Total		20

3.10.3 Cycle sequencing PCR program

The cycling conditions were

Initial denaturation:	96ºC for 1 minute;	
Denaturation:	96°C for 10 seconds;	25 cycles
Annealing:	50°C for 10 seconds;	
Extension:	60°C for 4 minutes;	
	Hold at 4°C until ready to purify	

3.10.4 Purification of the post cycle sequencing product

The Post cycle sequencing products were purified by ethanol precipitation to remove the unincorporated dye as follows.

- For each cycle sequencing reaction (20 µL) product, a 1.5 ml micro centrifuge tube was prepared containing the following
  - ✓ 2.0 µL of 3M sodium Acetate pH 4.6
  - ✓ 50.0 µL of 100% Ethanol
- The entire content of each extension reaction was pipetted into a tube of sodium acetate-ethanol-glycogen mixture and was mixed thoroughly.
- The tube was vortexed and was left at -30°C for 30 min to 1 hour to precipitate the extension product. (Notes: It should be kept in mind that longer extension period is not good as it can incorporate more salt)
- The tube was centrifuged at 13000 rpm for 20 min.
- The supernatant was carefully aspirated with a pipette tip and discarded.
- The precipitate was rinsed with 200 µL of 70% ethanol twice.
- The tube was vortexed briefly and was centrifuged for 10 min at 14000 rpm.
- The supernatant was carefully aspirated and the pellet was dried at R/T (40 min).
- The pellet was dissolved in 20 µL Hi-Di™ Formamide (16 µL, in case product size is <500bp).
- The solution was denatured at 95°C for 3 min in a thermal cycler and was kept in ice immediately.

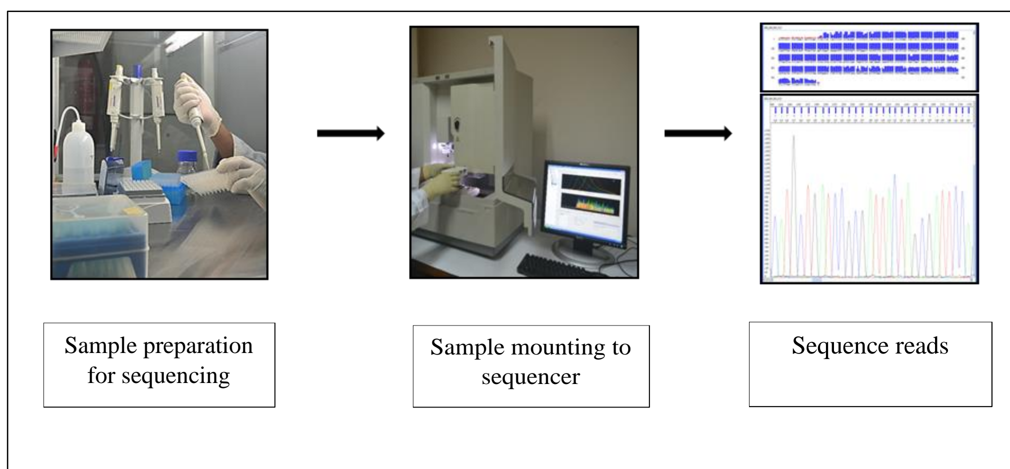


Figure 3.9: Steps of DNA sequencing processes.

3.10.5 Detection of the nucleotide by ABI PRISM® 3130 Genetic Analyzer

- After adding Hi-Di™ formamide, the sample was then provided for detection of the nucleotides by ABI PRISM® 3130 Genetic analyzer (**National Institute of Biotechnology, NIB**).
- DNA was separated through the pop 7 contained in the capillary and detected by the laser beam.
- When the nucleotide reached a detector window in capillary electrophoresis the fluorescent labelled fragments were excited by the laser beam of the machine.
- The laser excited fluorescent dye labels and emitted fluorescence was collected by the CCD camera. The fluorescence intensity data is interpreted into sequence data by specific software. The **green curve is for A (adenine)**, **blue for C (cytosine)**, **Red for T (thymine)** and **black is for G (guanine)**.

3.10.6 Sequence Data analysis

The chromatograms generated from the genetic analyzer along with the base sequences were analyzed by Bio edit Sequence Alignment editor.

3.11 Allele and genotype frequency calculation

The number of times the allele of interest is observed in a population is divided by the total number of copies of all the alleles at that particular genetic locus in the population to get an allele frequency. A decimal, a percentage, or a fraction can be used to denote

allele frequencies. The number of people with a certain genotype divided by the total number of people in a population is known as genotype frequency. The genotype frequency in population genetics is the frequency or proportion (i.e., $0 < f < 1$) of genotypes in a population.

3.12 Statistical Analyses

The normality of the random variables was confirmed by visual inspection of histograms using R statistical software version 4.0.3. The numerical variables of the GDM and control groups were compared using Student's t-tests, while the categorical variables were compared using the Chi-square test. The mean ± standard deviation of the mean (mean ± SD) was used to express numerical variables. The Shapiro-Wilk test was used to determine the normality of subgroup data. Categorical variables, on the other hand, were expressed as numbers (n) and percentages (%). When skewed distribution quantitative values were discovered, they were reported as median. The chi-square test and unpaired t-test were used to compare data in subgroups. The Mann-Whitney U test was used to quantify differences between groups for continuous variables having non-normal distributions. Statistical significance was defined as a P value ≤ 0.05.

Hardy-Weinberg equilibrium (HWE) was assessed in cases and controls separately using Pearson's chi-squared ($\chi2$) test with a $P>0.05$ criterion. Multivariate logistic regression was used to identify GDM risk variables. Using SNPStats, the general relationship of genotypes with GDM was examined using multivariate logistic regression analysis under codominant, dominant, recessive, overdominant, and log-additive models, and adjusted for family history of diabetes and gravidity (25). Using Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, the best-fitting model with many variables was chosen by stepwise inclusion of prospective confounding variables. SNPStats was used to investigate the correlation of target SNP haplotypes with GDM (25). The GAS power calculator (http://csg.sph.umich.edu/abecasis/cats/gas power calculator/index.html) was used to calculate statistical power.

3.13 Ethical consideration

Prior to the beginning of this study, the research protocol was approved by the research ethics committee (REC) of NIB (NIBREC 2016-04) (**Appendix A5**). Each screened patient enjoyed full rights either to participate or refuse or even withdraw from the study at any point of time. Written informed consent was taken from the participants. Proper medical services and advice was given to all subjects irrespective of status of enrollment. Information of the patient was kept confidential. Proper counseling was done before collection of blood samples. Adequate safety measure was also taken in every step of sample collection.

No drug or placebo was used for this study.

References

1. Kwak SH, Jang HC, Park KS. Finding genetic risk factors of gestational diabetes. Genomics Inform. 2012;10(4):239-43.

2. Galicia-Garcia U, Benito-Vicente A, Jebari S, Larrea-Sebal A, Siddiqi H, Uribe KB, et al. Pathophysiology of Type 2 Diabetes Mellitus. International journal of molecular sciences. 2020;21(17).

3. McCarthy MI. Genomics, Type 2 Diabetes, and Obesity. 2010;363(24):2339-50.

4. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007;445(7130):881-5.

5. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007;316(5829):1336-41.

6. El-Lebedy D, Ashmawy IJJoGE, Biotechnology. Common variants in TCF7L2 and CDKAL1 genes and risk of type 2 diabetes mellitus in Egyptians. 2016;14(2):247-51.

7. Han X, Luo Y, Ren Q, Zhang X, Wang F, Sun X, et al. Implication of genetic variants near SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, FTO, TCF2, KCNQ1, and WFS1 in type 2 diabetes in a Chinese population. 2010;11(1):1-9.

8. Bakhashab S, Filimban N, Altall RM, Nassir R, Qusti SY, Alqahtani MH, et al. The Effect Sizes of PPARγ rs1801282, FTO rs9939609, and MC4R rs2229616 Variants on Type 2 Diabetes Mellitus Risk among the Western Saudi Population: A Cross-Sectional Prospective Study. Genes (Basel). 2020;11(1):98.

9. Moniruzzaman M, Ahmed I, Huq S, All Mahmud MS, Begum S, Mahzabin Amin U, et al. Association of polymorphism in heat shock protein 70 genes with type 2 diabetes in Bangladeshi population. 2020;8(2):e1073.

10. Horikawa Y, Miyake K, Yasuda K, Enya M, Hirota Y, Yamagata K, et al. Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. The Journal of clinical endocrinology and metabolism. 2008;93(8):3136-41.

11. Tabara Y, Osawa H, Kawamoto R, Onuma H, Shimizu I, Miki T, et al. Replication study of candidate genes associated with type 2 diabetes based on genome-wide screening. Diabetes. 2009;58(2):493-8.

12. Wang J, Kuusisto J, Vänttinen M, Kuulasmaa T, Lindström J, Tuomilehto J, et al. Variants of transcription factor 7-like 2 (TCF7L2) gene predict conversion to type 2 diabetes in the Finnish Diabetes Prevention Study and are associated with impaired glucose regulation and impaired insulin secretion. Diabetologia. 2007;50(6):1192-200.

13. Scott LJ, Bonnycastle LL, Willer CJ, Sprau AG, Jackson AU, Narisu N, et al. Association of Transcription Factor 7-Like 2 (TCF7L2) Variants With Type 2 Diabetes in a Finnish Sample. Diabetes. 2006;55(9):2649-53.

14. Helgason A, Pálsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, et al. Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. Nature genetics. 2007;39(2):218-25.

15. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. The Journal of clinical investigation. 2007;117(8):2155-63.

16. Li T, Xu X, Xu Y, Jin P, Chen J, Shi Y, et al. PPARG Polymorphisms Are Associated with Unexplained Mild Vision Loss in Patients with Type 2 Diabetes Mellitus. Journal of ophthalmology. 2019;2019:5284867.

17. Lv X, Zhang L, Sun J, Cai Z, Gu Q, Zhang R, et al. Interaction between peroxisome proliferator-activated receptor gamma polymorphism and obesity on type 2 diabetes in a Chinese Han population. Diabetol Metab Syndr. 2017;9:7.

18. Neuman RJ, Wasson J, Atzmon G, Wainstein J, Yerushalmi Y, Cohen J, et al. Gene-gene interactions lead to higher risk for development of type 2 diabetes in an Ashkenazi Jewish population. PloS one. 2010;5(3):e9903.

19. Chen G, Xu Y, Lin Y, Lai X, Yao J, Huang B, et al. Association study of genetic variants of 17 diabetes-related genes/loci and cardiovascular risk and diabetic nephropathy in the Chinese She population (中国畲族人群17个糖尿病相关基因位点的遗传变异与心血管风险和糖尿病肾病的相关性). 2013;5(2):136-45.

20. Xiao S, Zeng X, Quan L, Zhu J. Correlation between polymorphism of FTO gene and type 2 diabetes mellitus in Uygur people from northwest China. Int J Clin Exp Med. 2015;8(6):9744-50.

21. Sokup A, Ruszkowska-Ciastek B, Góralczyk K, Walentowicz M, Szymański M, Rość D. Insulin resistance as estimated by the homeostatic method at diagnosis of gestational diabetes: estimation of disease severity and therapeutic needs in a population-based study. BMC endocrine disorders. 2013;13:21.

22. Russell JFSaDW. Molecular Cloning: A Laboratory Manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.

23. Santos FR, Bianchi NO, Pena SJGR. Worldwide distribution of human Y-chromosome haplotypes. 1996;6(7):601-11.

24. De la Vega FM, Lazaruk KD, Rhodes MD, Wenz MH. Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP Genotyping Assays and the SNPlex Genotyping System. Mutation research. 2005;573(1-2):111-35.

25. Solé X, Guinó E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. Bioinformatics (Oxford, England). 2006;22(15):1928-9.

4. Frequency Detection of the target SNPs related to T2D in women with and without GDM

4.1 Study subjects

A total of 534 pregnant women were screened for having any degree of hyperglycemia, irrespective of the trimester. According to exclusion criteria, 29 women were excluded from the study, among whom 17 were diagnosed with DIP (**Section 2.3 and Table 3.3**), three were recipients of steroid treatment, five had thyroid dysfunction, and four did not repeat OGTT after the 24th week of gestation (**Section 3.2.2**). Finally, 505 samples were selected, among which 286 were normoglycemic and 219 were diagnosed with gestational diabetes mellitus.

4.1.1 General characteristics of the study subjects

Differences in the general characteristics of 219 pregnant women with GDM and 286 control subjects are shown in the following table (**Table 4.1**). They were 18 to 44 years old (years). GDM patients were substantially older, had a higher BMI, had higher diastolic blood pressure, and had higher plasma glucose levels than controls. The percentage of the positive family history of diabetes and multigravida was significantly higher in the GDM group. The percentage of primigravida was significantly higher in the normoglycemic control group. There were no significant differences between systolic blood pressure, percentage of pregnancies with bad obstetric history, occupation, and maternal history of diabetes between these two groups.

Table 4.1: General characteristics of the study participants

Variables	Control(n=286)	GDM(n=219)	*P* value
Age, years	25.47±4.77	27.57±4.59	**<0.0001**
BMI (kg/m^2)	25.19±3.83	26.70±4.09	**<0.0001**
SBP (mmHg)	108.66±11.98	109.26±11.77	0.5744
DBP (mmHg)	68.83±8.96	70.65±9.14	**0.03**
Plasma glucose levels(mmol/L)			
FPG	4.33±0.46	5.15±0.69	**< 0.0001**
OPG	7.50±1.22	9.88±1.58	**< 0.0001**
TPG	6.43±1.05	8.27±1.48	**< 0.0001**
Family history of Diabetes, number (%)			
No	194(67.83%)	117(53.42%)	**0.0010**
Yes	92(32.17%)	102 (46.58%)	
a. **Father**	18(6.29%)	27(12.33%)	**0.0183**
b. **Mother**	35(12.24%)	26(11.87%)	0.8995
c. **Both**	35(12.24%)	42(19.18%)	**0.0317**
d. **Siblings**	4(1.39%)	7(3.19%)	0.1694
Gravidity, number (%)			
a. **Primigravida**	133(46.50%)	81(36.99%)	**0.03**
b. **Multigravida**	153(53.50%)	138(63.01%)	
Bad obstetric history, number (%)			
Yes	74(25.87%)	59(26.94%)	0.787
No	212(74.13%)	160(73.06%)	
Occupation, number (%)			
a. **Housewife**	190(66.43%)	148(67.58%)	0.786
b. **Service Holder**	58(20.28%)	56(25.57%)	0.159
c. **Others**	38(13.29%)	15(6.85%)	**0.019**

Data are presented either as mean ± standard deviation or number with percent in parenthesis. *P*-values are estimated by independent sample t-test or chi-square test. BMI, Body mass index. SBP, systolic blood pressure. DBP, diastolic blood pressure. FPG, fasting plasma glucose. OPG, plasma glucose level after one hour of the glucose load. TPG, plasma glucose level after two-hour of the glucose load.

4.2 Frequency Detection of the Type 2 Diabetes (T2D) Related target SNPs

Genome-wide association studies (GWAS) revealed the association of a number of genes with T2DM (1-5). Replication of these associations in different populations and the meta-analysis for each of these genes also confirmed the risk for developing T2DM (6-8). Genetic investigations have also discovered a preliminary set of T2D-associated loci employing linkage analysis and candidate gene approaches (9-12). This study examined the genotype and allele frequencies of the selected SNPs of six genes in the participants (**Section 3.1.1 and 3.1.2**). Hardy-Weinberg Equilibrium (HWE) is used for genotype frequency estimation of variants based on its allele frequency in non-evolving populations (13). After affirming that the resulting frequencies were in agreement with HWE, further analyses were carried out to determine their association with GDM.

4.2.1 Frequency detection of the *CDKAL1* gene variants rs7756992 and rs7754840

4.2.1.1 General characteristics of the subjects genotyped for the polymorphisms

CDKAL1 gene variants rs7756992 and rs7754840 were studied in 468 participants in total, with 212 women with GDM and 256 healthy controls.

Table 4.2: Anthropometric and demographic data of the participants(n=468)

	Controls n=256	**GDM n=212**	***P* value**
Age	25.42±4.58	27.58±4.59	**< 0.0001**
BMI	25.29±3.92	26.64±4.15	**0.0003**
FPG	4.31±0.48	5.15±0.47	**<0.0001**
OPG	7.49±1.26	9.77±1.65	**<0.0001**
TPG	6.4±1.06	8.25±1.49	**< 0.0001**
Positive Family History of Diabetes	79(30.85%)	98(46.23%)	**0.0006**
Multigravida	137(53.52%)	133(62.74%)	**0.04**
Primigravida	119(46.48%)	79(37.26%)	**0.04**
SBP	109.27±11.75	108.94±11.91	0.69
DBP	70.51±9.18	69.25±9.05	0.12

Data are presented either as mean ± standard deviation or as percentage. *P*-values are estimated by independent sample t-test or chi-square test. BMI, Body mass index. SBP, systolic blood pressure. DBP, diastolic blood pressure. FPG, fasting plasma glucose. OPG, plasma glucose level after one hour of the glucose load. TPG, plasma glucose level after two-hour of the glucose load,

4.2.1.2 Genotype and allele frequencies of the *CDKAL1* gene variants

The *CDKAL1* gene variant rs7756992 was genotyped by PCR-RFLP [Figure 4.1(a)]. The AA genotype was revealed by the development of two bands of 388 and 296 bp, whereas the GG genotype was indicated by the 684 bp digestion product. All three bands were visible when heterozygous genotypes were present [Figure 4.1(a)]. Genotypes obtained were confirmed by sequencing the blindly selected samples [Figure 4.1(b)]. After sequencing, no changes in genotypes were found.

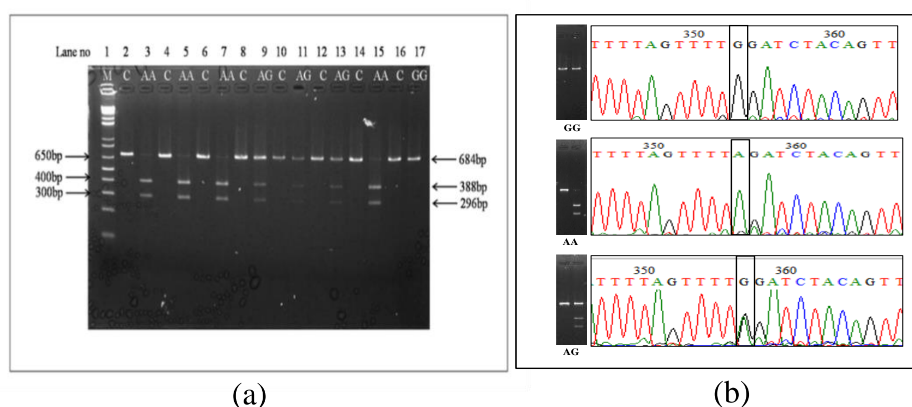


Figure 4.1: (a) RFLP analysis for genotyping. Lane 1(M): 1kb plus DNA ladder, lanes 2,4,6,8,10,12,14 and 16; Undigested PCR product (C=control); lane 3, 5, 7 and 15; AA Homozygous genotype, lanes 9, 11 and 13; AG Heterozygous genotypes and lane 17; GG Homozygous genotype. (b) Confirmation of the RFLP results by DNA sequencing.

The second variant, rs7754840, of this gene was genotyped by TaqMan^{TM} allele discrimination assay (Figure 4.2). The repetition of the assay was carried out in 20% of the samples to confirm the resulted genotypes. After repeating the assay on blindly selected samples, no changes in the genotypes were found.

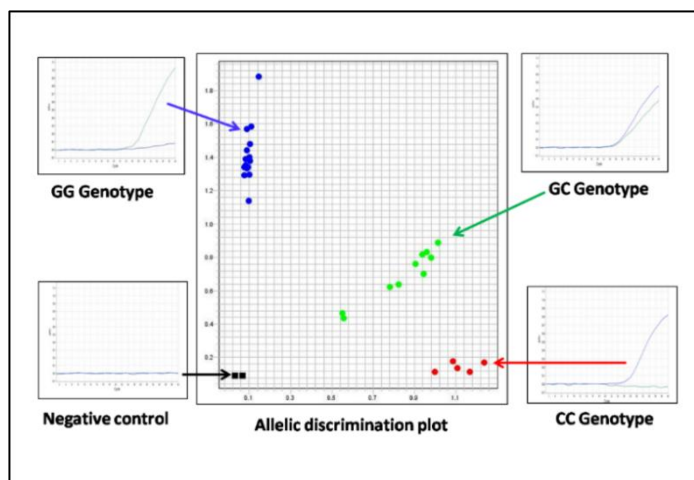


Figure 4.2: A single run of 30 samples yielded an allele discrimination plot that comprised both cases and controls, as well as representative amplification plots for each genotype and a negative control. Negative controls were represented by black squares, while GG, GC, and CC genotypes were represented by blue, green, and red dots, respectively.

The genotype distribution of the *CDKAL1* gene polymorphisms rs7756992 and rs7754840 demonstrated that these two SNPs differed between participants with and without GDM (Table 4.3). The heterozygote genotype (AG) of rs7756992 and homozygotes of altered alleles (GG of rs7756992 and CC of rs7754840) were notably higher in the GDM group. When the frequency of the genotypes containing the risk allele is higher in the disease group, it indicates the risk implementing nature of the SNP. The genotype distributions of the case and control groups of rs7754840 and the control group of rs7756992 were consistent with Hardy-Weinberg equilibrium (HWE) (Table 4.3). The genotype distribution of **rs7756992 in cases departed** from HWE. The frequencies of the risk alleles of both SNPs were higher in GDM group.

Table 4.3: The genotype and allele frequency of *CDKAL1* gene variants (rs7756992 and rs7754840) in study participants

SNP	Genotype/Allele	Control (%)	GDM (%)
rs7756992	AA	123 (48%)	78 (36.8%)
	AG	114 (44.5%)	113 (53.3%)
	GG	19 (7.4%)	21 (9.9%)
	A	360 (70.31%)	269(63.44%)
	G	153(29.65%)	155(36.56%)

HWE	Chi-Square (χ^2)	1.14	4.71
	P value	0.29	**0.03**
rs7754840	GG	141 (55.1%)	107 (50.5%)
	GC	102 (39.8%)	85 (40.1%)
	CC	13 (5.1%)	20 (9.4%)
	G	384 (75%)	299(71%)
	C	128 (25%)	125(29%)
HWE	Chi-Square (χ^2)	1	0.27
	P value	0.317	0.603

If $P < 0.05$ - not consistent with Hardy Weinberg Equilibrium (HWE).

4.2.2 Frequency detection of the *FTO* gene variant rs8050136

The *FTO* gene variant rs8050136 was investigated in 502 pregnant women, 218 of whom had GDM and 284 of whom were not. The T-ARMS-PCR technique was used to genotype this variation (**Section 3.9.3**). The AA genotype was represented by two bands of 337 bp and 251 bp, while the CC genotype was represented by 337 bp and 130 bp products. Three bands of 337, 251, and 130 bp were also found when heterozygous genotype AC was present [Figure 4.3 (a)]. The resulting genotypes were confirmed by sequencing blindly selected samples [Figure 4.3 (b)]. No difference in the genotypes was obtained after sequencing.

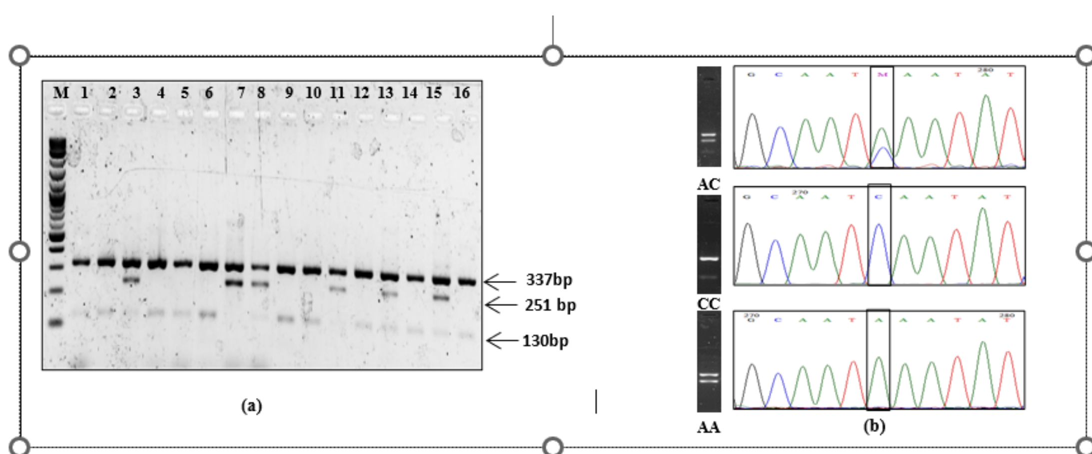


Figure 4.3: (a) Lane1 (M): 1Kb+ DNA Ladder. Lane 1, 2, 4, 5, 6, 9, 10, 12, 14, 16 homozygous for wild type (CC). Lane 3, 8, 13, and 15 are heterozygous (AC). Lane 7 and 11 are homozygous for mutant allele (AA). (b)Validation of the T-ARMS results by DNA sequencing.

Genotyping for the FTO gene polymorphism showed that the genotype distribution of this SNP differed between those with and without GDM (Table 4.4). The frequencies of risk (AC and AA) genotypes of this SNP were higher in GDM group, suggesting the risk providing nature of this polymorphism. Genotype distributions of cases (GDM) and controls were consistent with Hardy-Weinberg equilibrium (HWE) (Table 4.4). The frequencies of risk (AC and AA) genotypes of this SNP were higher in GDM group, suggesting the risk providing nature of this polymorphism. Genotype distributions of cases (GDM) and controls were consistent with Hardy-Weinberg equilibrium (HWE) (Table 4.4). The minor allele (A) frequency was higher in GDM group.

Table 4.4: Genotype and allele frequency of *FTO* gene variant rs8050136 in the study participants

SNP	Genotype/Allele	Control (%)	GDM (%)
rs8050136	CC	143 (50.35%)	101 (46.33%)
	AC	126 (44.37%)	98 (44.95%)
	AA	15 (5.28%)	19 (8.72%)
	C	412(72.54%)	300(68.81%)
	A	156(27.46%)	136(31.19%)
HWE	Chi-Square (χ^2)	1.29	0.22
	P value	0.26	0.64

If P < 0.05 - not consistent with Hardy Weinberg Equilibrium (HWE).

4.2.3 Frequency detection of the *HSPA1L* gene variant rs2227956

A total of 501 pregnant women comprising 218 with GDM and 283 normoglycemic control subject were genotyped for frequency detection of *HSPA1L* gene variant rs2227956. The variant was genotyped by the PCR-RFLP method (Section 3.8.1). The TT genotype was indicated by the development of two bands of 616 bp and 246 bp, whereas the CC genotype was indicated by 862 bp digestion products. All three bands were visible when heterozygous genotypes were present [Figure 4.4(a)]. The genotypes were confirmed by sequencing samples at random [Figure 4.4(b)].

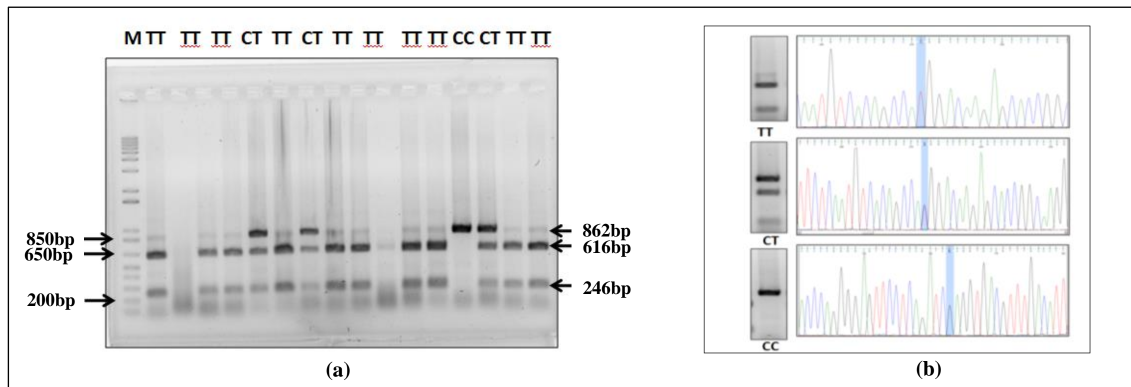


Figure 4.4: (a) Lane 1(M): 1kb plus DNA ladder, lanes 2,4,5,7,9,11,12,15 and 16: TT homozygous; lanes 6,8 and 14: CT heterozygous; lane 13 CC homozygous genotypes. (b) Confirmation of rs2227956 genotypes by DNA sequencing

The allele and genotype frequencies of this SNP in cases and controls are shown in Table 4.5. The distribution of the risk genotype CC was low both in control and cases (GDM). The heterozygote (CT) genotype frequency was higher in the control group, whereas the homozygous risk genotype (CC) was higher in the GDM group. Genotype distributions of cases and controls were consistent with HWE (Table 4.5). The frequencies of the minor allele C of this SNP are identical in both groups. These discrepancies in these two groups' genotypes and allele frequencies suggested a lack of risk-providing nature of this SNP.

Table 4.5: Genotype and allele frequency of *HSPA1L* gene variant rs2227956 in the study participants

SNP	Genotype/Allele	Control (%)	GDM (%)
rs2227956	TT	210 (74.2%)	162 (74.3%)
	CT	66 (23.3%)	50 (22.9%)
	CC	7 (2.5%)	6 (2.8%)
	T	486(86%)	374(86%)
	C	**80(14%)**	**62(14%)**
HWE	Chi-Square (χ^2)	0.43	0.78
	P value	0.51	0.38

If P < 0.05 - not consistent with Hardy Weinberg Equilibrium (HWE).

4.2.4 Frequency detection of the *PPARG* gene variants rs3856806 and rs1801282

The target SNP of the *PPARG* gene, rs3856806, was genotyped in 502 participants, consisting of 285 normoglycemic controls and 217 GDM cases. On the other hand, the rs1801282 polymorphism of the same gene was genotyped in 505 participants, comprising 286 controls and 219 GDM cases.

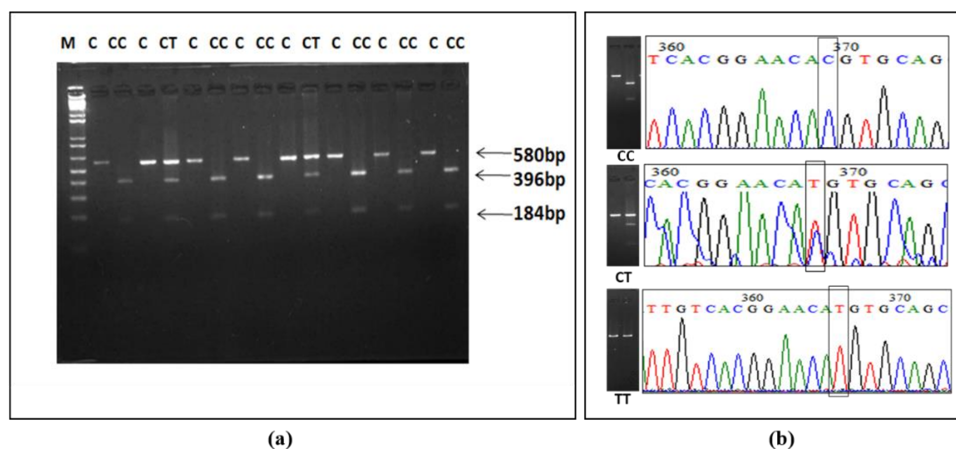


Figure 4.5: (a) Lane 1 (M): 1kb plus DNA ladder, lanes 2,4,6,8,10,12,14, and 16; (C=Control) PCR product(undigested); lane 3,7,9,13,15 and 17; CC Homozygous, lanes 5 and 11; CT genotypes. (b) Confirmation of rs3856806 genotypes by DNA sequencing

The variants, rs3856806 and rs1801282, of the *PPARG* gene were genotyped by PCR-RFLP [Figure 4.5(a)] and TaqMan allele discrimination assay (Figure 4.6), respectively. The CC genotype was indicated by the presence of two bands of 396 bp and 184 bp, whereas the TT genotype was indicated by 580 bp digestion products. All three bands were visible when heterozygous genotypes were present [Figure 4.5(a)]. Sequencing selected samples confirmed genotypes obtained by the PCR-RFLP method, and those by TaqMan assay were confirmed by repeating the assay in 20% of the total samples.

The allele and genotype frequencies of these two SNPs in cases and controls are shown in Table 4.6. The frequencies of the heterozygotes (CT and CG) of the two SNPs and homozygote (GG) of rs1801282 of this gene were found to be higher in controls (non-GDM) than in the cases (GDM), suggesting the probable absence of risk manipulating nature of the altered allele.

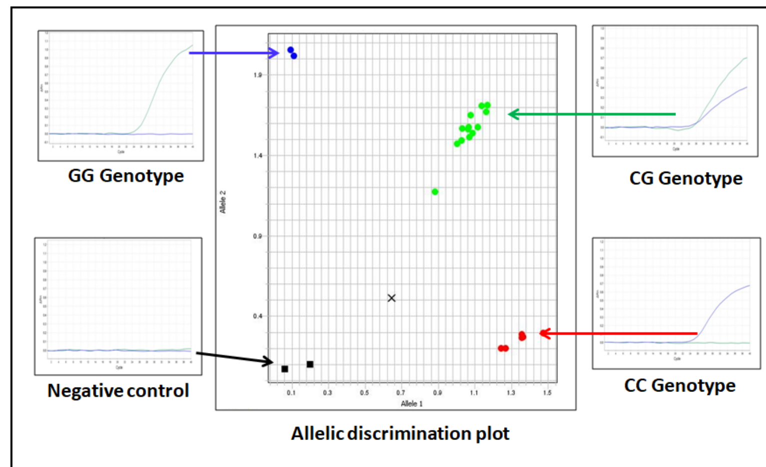


Figure 4.6: An allele discrimination plot generated from a single run of 30 samples, which included both cases and controls, as well as representative amplification plots for each genotype and a negative control. GG, GC, and CC genotypes were represented by blue, green, and red dots, respectively; negative controls were represented by black squares.

A similar combination pattern was not observed in the homozygote genotype frequency (TT) of rs3856806. The homozygotes (TT and GG) of the altered allele of both SNPs have been found to have the least frequency. The genotype frequencies of the two SNPs were all in Hardy-Weinberg equilibrium among the controls and cases (all P>0.05). The presence of the risk alleles of *PPARG* SNPs, T allele of rs3856806, and G allele of rs1801282 was relatively not higher in GDM patients than in the NGT group. The frequency of the T allele of rs3856806 is identical in both groups (Table 4.6).

Table 4.6: Genotype and allele frequency of *PPARG* gene variants (rs3856806 and rs1801282) in the study participants

SNP	Genotype/Allele	Control (%)	GDM (%)
rs3856806	CC	194(68.1%)	148(68.2%)
	CT	81(28.4%)	61(28.1%)
	TT	10(3.5%)	8(3.7%)
	C	469(82%)	357(82%)
	T	101(18%)	77(18%)
HWE	Chi-Square (χ^2)	0.18	0.295
	P value	0.67	0.59

rs1801282	CC	213(75%)	170(78%)
	CG	68(24%)	46(21%)
	GG	5(2%)	3(1%)
	C	494(86%)	386(88%)
	G	78(14%)	52(12%)
HWE	Chi-Square (χ^2)	0.026	0.003
	P value	0.873	0.955

If P < 0.05 - not consistent with Hardy Weinberg Equilibrium (HWE).

4.2.5 Frequency detection of the *TCF7L2* gene variants, rs12255372, rs10885406 and rs7903146

The target SNP of the *TCF7L2* gene, rs10885406, was genotyped in all 505 participants, consisting of 286 normoglycemic controls and 219 GDM cases. On the other hand, the rs7903146 and rs12255372 polymorphisms of the same gene were genotyped in 501 participants, comprising 284 controls and 217 GDM cases, and 502 participants, comprising 285 controls and 217 GDM cases, respectively.

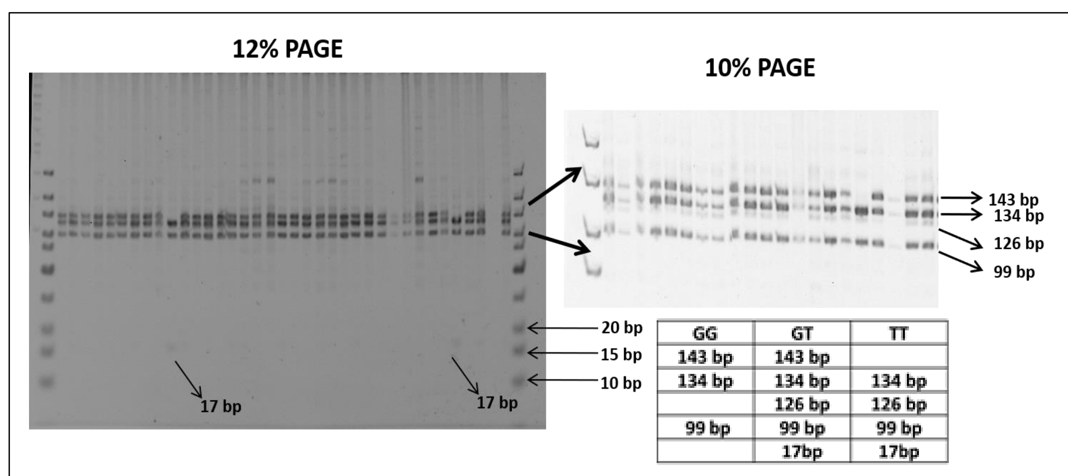


Figure 4.7: Lane 1 and 40 (M): 1kb plus DNA ladder; lanes 2-39 digested PCR products for rs12255372 genotyping

One of the variants, rs12255372, was genotyped by the PCR-RFLP method (Figure 4.7). The lengths of digestion products and their corresponding genotypes are shown in figure 4.7. Genotypes obtained were confirmed by sequencing blindly selected samples (Figure 4.8).

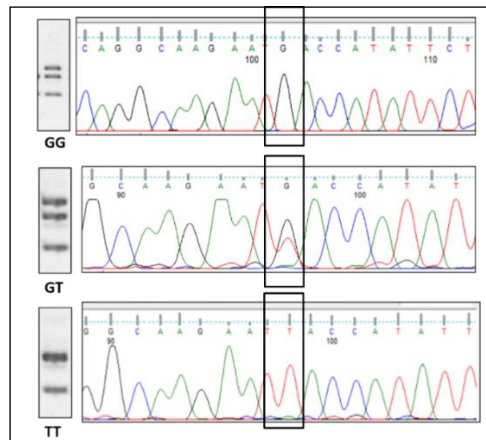


Figure 4.8: Confirmation of the rs12255372 of *TCF7L2* gene by sequencing blindly selected samples

The *TCF7L2* gene variants, rs10885406 and rs7903146, were genotyped (Figure 4.9) by the allelic discrimination assay (Section 3.9.2). Genotypes obtained by TaqMan assay were confirmed by repeating the assay in 20% of the samples.

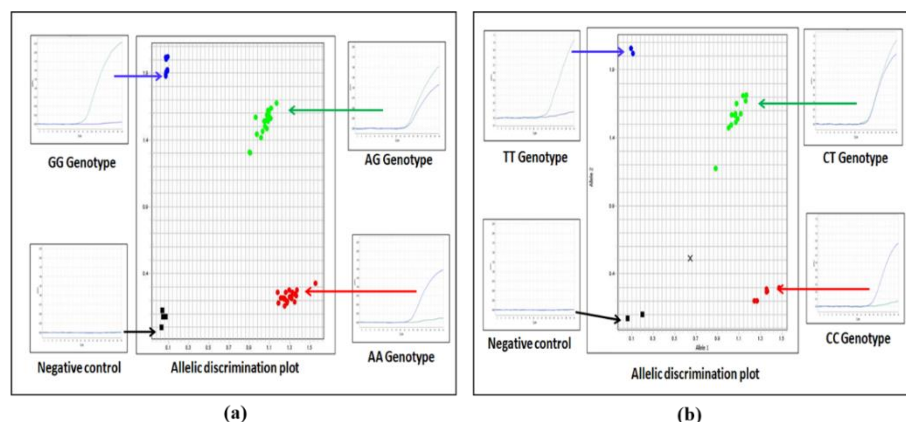


Figure 4.9: Allele discrimination plots (a) rs10885406 and (b) rs7903146 from two single runs of 35 samples containing both controls and cases and amplification plots for each genotype, as well as a negative control. Blue, green and red dots denoted in (a) GG, AG and AA genotypes; and in (b) TT, CT and CC genotypes respectively; black squares showed negative controls in both plots.

The allele and genotype frequencies of rs10885406 in cases and controls are shown in Table 4.7. Genotyping of this variant revealed the genotypes differed in participants with and without GDM except that the frequency of AA genotype was equal between these two groups. The higher frequency of the homozygous risk genotype (GG) and that of the G allele in the control group, suggested the protective nature of this SNP. Genotype distributions of cases (GDM) and controls were consistent with HWE (Table 4.7).

Table 4.7: Genotype and allele frequency of *TCF7L2* gene variant rs10885406 in the study participants

SNP	Genotype/Allele	Control (%)	GDM (%)
rs10885406	AA	111 (38.8%)	85 (38.8%)
	AG	140 (49%)	111 (50.7%)
	GG	35 (12.2%)	23 (10.5%)
	A	362(63%)	281(64%)
	G	210(37%)	157(36%)
HWE	Chi-Square (χ^2)	0.816	2.28
	P value	0.37	0.131

If P < 0.05 - not consistent with Hardy Weinberg Equilibrium (HWE).

The allele and genotype frequencies of rs7903146 in cases and controls are shown in Table 4.8. The frequency of the heterozygote (CT) of the risk allele was almost equal in both groups, whereas the homozygote (TT) was higher in the control group. The risk allele (T) frequency was also higher in the control group. These discrepancies in genotype and allele frequencies indicate the protective nature of this polymorphism. Genotype distributions of cases (GDM) and controls were consistent with HWE.

Table 4.8: Genotype and allele frequency of the *TCF7L2* gene variant rs7903146 in the study subjects

SNP	Genotype/Allele	Control (%)	GDM (%)
rs7903146	CC	142 (50%)	112 (51.6%)
	CT	122 (43%)	93 (42.9%)
	TT	20 (7%)	12 (5.5%)
	C	406(71%)	317(73%)
	T	162(29%)	117(27%)
HWE	Chi-Square (χ^2)	0.815	1.689
	P value	0.37	0.19

If P < 0.05 - not consistent with Hardy Weinberg Equilibrium (HWE).

The allele and genotype frequencies of rs12255372 in cases and controls are shown in Table 4.9. Genotyping revealed that genotypes of this SNP differed in individuals with

and without GDM. The frequencies of heterozygote (GT) and homozygote (TT) genotypes and the altered allele T were higher in the GDM group, suggesting the risk manipulating nature of this SNP. The genotypes of the control group were consistent with HWE (Table 4.9), whereas in GDM group, the genotype distribution was departed from HWE.

Table 4.9: Genotype and allele frequency of *TCF7L2* gene variant rs12255372 in the study subjects

SNP	Genotype/Allele	Control (%)	GDM (%)
rs12255372	GG	154 (54%)	95 (43.8%)
	GT	118 (41.4%)	111 (51.1%)
	TT	13 (4.6%)	11 (5.1%)
	G	426(75%)	301(69%)
	T	144(25%)	133(31%)
HWE	Chi-Square (χ^2)	2.65	8.97
	P value	0.103	**0.002**

If P < 0.05 - not consistent with Hardy Weinberg Equilibrium (HWE).

4.2.6 Frequency detection of the *WFS1* gene variant, rs10010131

The selected SNP of the *WFS1* gene was genotyped by the PCR-RFLP method [Figure 4.10 (a)] in 502 participants of this study. Among them, 283 were normoglycemic control, and 219 were GDM cases.

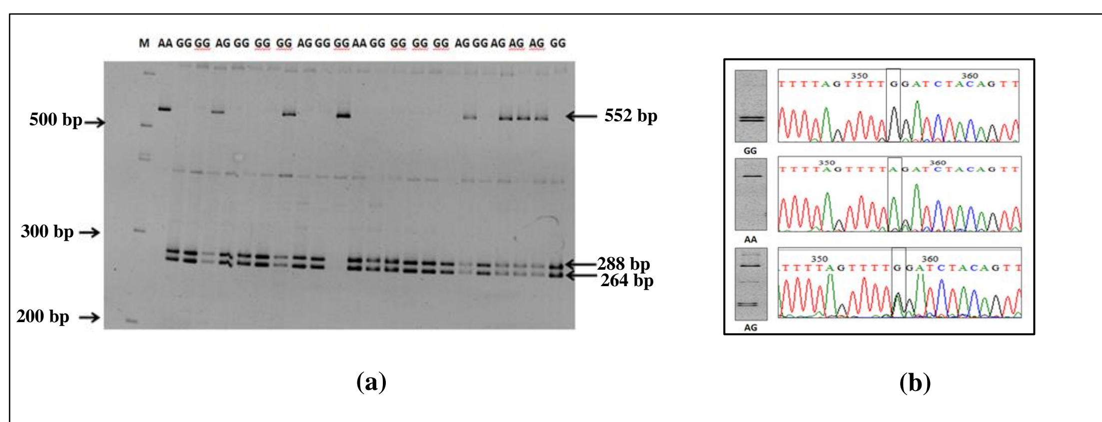


Figure 4.10: (a) Lane 1(M): 1kb plus DNA ladder, lanes 2 and 12: AA homozygous; lanes 3,4,6,7,8,10,11,13-18,20 and 24: GG homozygous; lanes 5,9,19 and 21-23: AG heterozygous genotypes. (b) Confirmation of the RFLP results by DNA sequencing.

Genotyping for rs10010131 polymorphism of the *WFS1* gene revealed that genotype distribution of this SNP differed in cases and controls (Table 4.8). The frequency of the heterozygote (AG) of this variant was found to be higher in controls (non-GDM) than in the cases (GDM), suggesting the probable absence of risk manipulating nature of the altered allele. But the homozygotes of the risk allele (AA) and reference allele (GG) was found higher in frequency in GDM group. The genotype frequencies of this SNP were all in Hardy-Weinberg equilibrium among the controls and cases (both $P>0.05$).

Table 4.10: Genotype and allele frequency of the *WFS1* gene variant rs10010131 in the study participants

SNP	Genotype/Allele	Control (%)	GDM (%)
rs10010131	GG	153 (54.1%)	124 (56.6%)
	AG	113 (39.9%)	76(34.7%)
	AA	17(6%)	19(8.7%)
	G	419(74.03%)	324(73.97%)
	A	147(25.97%)	114(26.03%)
HWE	Chi-Square (χ^2)	0.417	2.14
	P value	0.52	0.14

If $P < 0.05$ - not consistent with Hardy Weinberg Equilibrium (HWE).

4.3 Expected and observed genotype and allele frequencies of selected SNPs in the study participants

The observed and expected frequencies of genotypes and alleles in controls and cases are shown in figure 4.11 and figure 4.12, respectively.

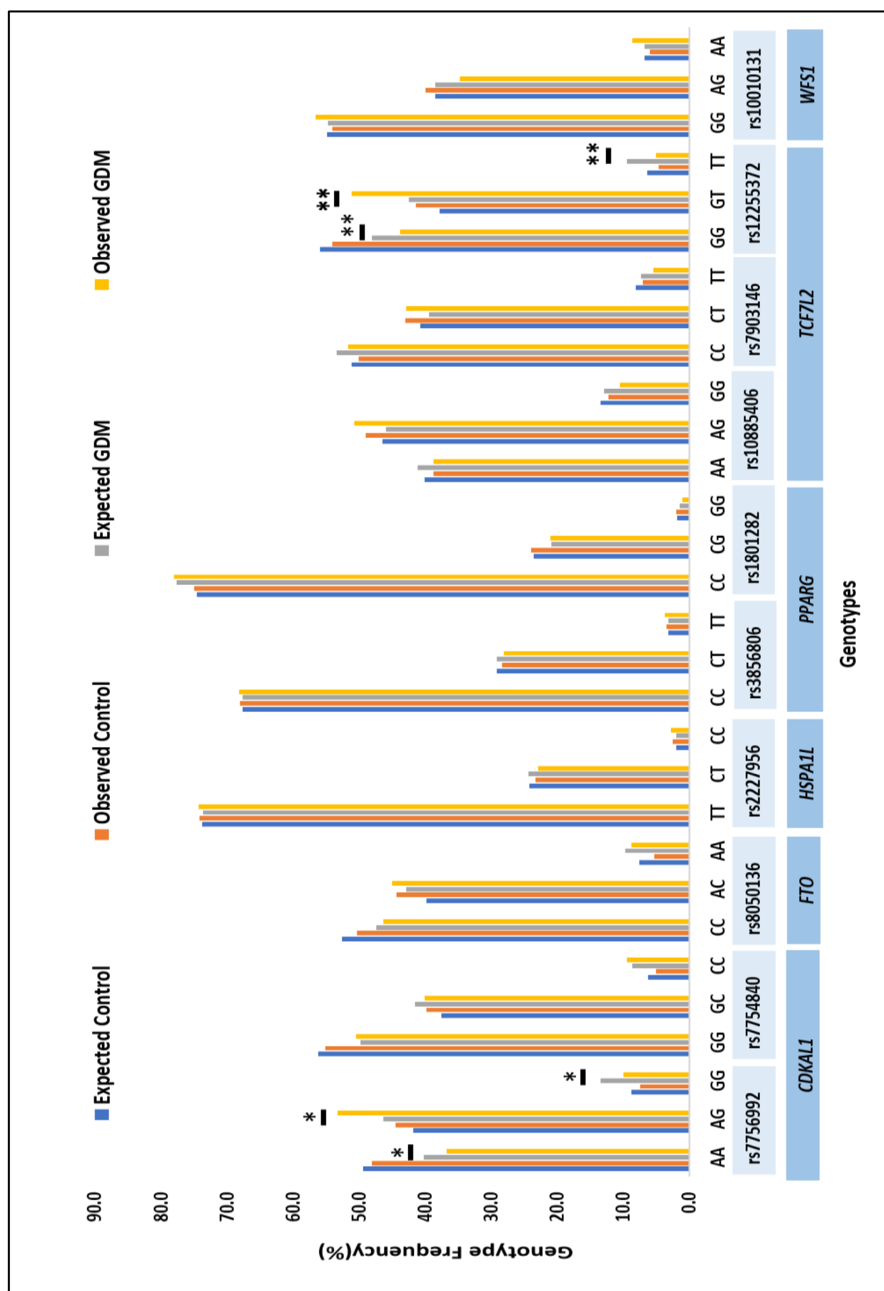


Figure 4.11: Comparison of expected and observed genotype frequencies of the target SNPs in control and cases.

Significant differences between observed and expected frequencies were observed in cases of the *CDKAL1* gene variant rs7756992 and the *TCF7L2* gene variant rs12255372. There was no difference obtained between expected and observed allele frequencies in cases and controls (Figure 4.12).

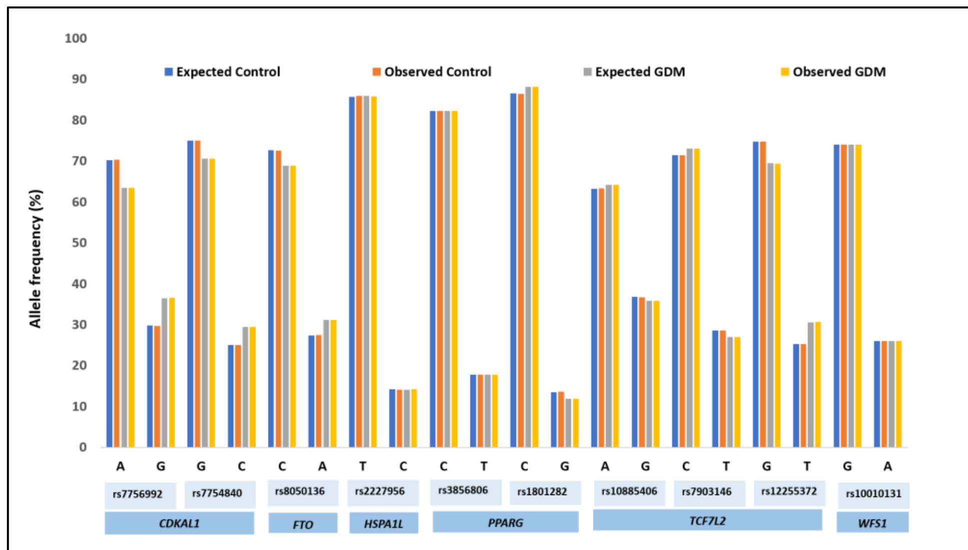


Figure 4.12: Comparison of expected and observed allele frequencies of the target SNPs in control and cases.

4.4 The frequency detection of the target SNPs in the study participants

The distribution of genotypes and alleles of the target SNPs in the total study participants are shown in the table 4.9.

Table 4.11: Genotype and allele frequency of the target SNPs in the study participants(n=505)

Gene	SNP Id	Genotype	Frequency (%)	Allele	Frequency
CDKAL1	rs7756992	AA	42.95	A	0.672
		AG	48.50	**G**	0.328
		GG	8.55		
	rs7754840	GG	52.99	G	0.730
		GC	39.96	**C**	0.270
		CC	7.05		
FTO	rs8050136	CC	48.61	C	0.709
		AC	44.62	**A**	0.291
		AA	6.77		
HSPA1L	rs2227956	TT	74.25	T	0.86
		CT	23.15	**C**	0.14
		CC	2.59		
PPARG	rs3856806	CC	68.13	C	0.82
		CT	28.29	**T**	0.18
		TT	3.59		
	rs1801282	CC	75.84	C	0.87
		CG	22.57	**G**	0.13
		GG	1.58		
TCF7L2	rs10885406	AA	38.81	A	0.64
		AG	49.70	**G**	0.36

		GG	11.49		
	rs7903146	CC	50.70	C	0.72
		CT	42.91	**T**	0.28
		TT	6.39		
	rs12255372	GG	49.60	G	0.72
		GT	45.62	**T**	0.28
		TT	4.78		
WFS1	rs10010131	GG	55.18	G	0.74
		AG	37.65	**A**	0.26
		AA	7.17		

4.5 Summary

- Sample size of the study is 505 comprising 286 normoglycemic controls and 219 GDM cases.
- Genotype frequencies of the selected 10 SNPs in the control group were in agreement with HWE.
- In the GDM group, genotype frequencies of eight SNPs were in accordance with HWE.
- The *CDKAL1* gene variant rs7756992 and the *TCF7L2* gene variant rs12255372 were not in agreement ($P<0.05$) with HWE in the GDM group.
- The genotype and allele frequencies of the following SNPs indicate the possible risk manipulating nature of these variants:

Sl. No	Gene	SNP
1	*CDKAL1*	rs7756992
2	*CDKAL1*	rs7754840
3	*FTO*	rs8050136
4	*TCF7L2*	rs12255372
5	*WFS1*	rs10010131

- The genotype and allele frequencies of the following SNPs indicate the possible protective nature of these variants:

Sl. No	Gene	SNP
1	*HSPA1L*	rs2227956
2	*PPARG*	rs1801282
3	*TCF7L2*	rs10885406
4	*TCF7L2*	rs7903146

- The genotype and allele frequencies of *PPARG* rs3856806 were equal in both groups and indicates no association with GDM.

References

1. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007;445(7130):881-5.

2. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007;316(5829):1336-41.

3. Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, et al. Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. Diabetes. 2009;58(2):505-10.

4. Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, Ohnaka K, et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. Diabetes. 2009;58(7):1690-9.

5. Chen J, Sun M, Adeyemo A, Pirie F, Carstensen T, Pomilla C, et al. Genome-wide association study of type 2 diabetes in Africa. 2019;62(7):1204-11.

6. Phani NM, Adhikari P, Nagri SK, D'Souza SC, Satyamoorthy K, Rai PS. Replication and Relevance of Multiple Susceptibility Loci Discovered from Genome Wide Association Studies for Type 2 Diabetes in an Indian Population. PloS one. 2016;11(6):e0157364.

7. Parra E, Below J, Krithika S, Valladares A, Barta J, Cox N, et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. 2011;54(8):2038-46.

8. De Cosmo S, Prudente S, Lamacchia O, Lapice E, Morini E, Di Paola R, et al. PPARγ2 P12A polymorphism and albuminuria in patients with type 2 diabetes: a meta-analysis of case–control studies. 2011;26(12):4011-6.

9. Khetan S, Kursawe R, Youn A, Lawlor N, Jillette A, Marquez EJ, et al. Type 2 Diabetes-Associated Genetic Variants Regulate Chromatin Accessibility in Human Islets. Diabetes. 2018;67(11):2466-77.

10. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes. 2017;66(11):2888-902.

11. Wu L, Wang CC. Genetic variants in promoter regions associated with type 2 diabetes mellitus: A large-scale meta-analysis and subgroup analysis. Journal of cellular biochemistry. 2019;120(8):13012-25.

12. Xue A, Wu Y, Zhu Z, Zhang F, Kemper KE, Zheng Z, et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nature communications. 2018;9(1):2941.

13. Abramovs N, Brass A, Tassabehji M. Hardy-Weinberg Equilibrium in the Large Scale Genomic Sequencing Era. 2020;11.

5. Association of T2DM related SNPs with the predisposition of GDM

5.1 Association of T2DM linked genetic variants with GDM

In this case-control study, the prevalence of GDM has been compared among individuals with normal alleles and individuals with variant alleles that result in an odds ratio (OR). Odds are given for each genotype of an SNP, and a pair of odds generates an OR. For further confirmation whether each SNP (**Table 3.2**) is associated with GDM, whether the probability of this disease increases with the increase of the number of risk alleles in the genotype, the Cochran-Armitage trend test has been carried out under five different genetic models (codominant model, dominant model, recessive model, overdominant and log-additive model). Confounding effects of the studied variables (**Table 4.1**) on association have been detected. Among the studied variables age, BMI, DBP, occupation (others), gravidity, and family history of diabetes (FHD) were significantly different in the control and GDM groups and were the prospective confounders. Association of these variables with GDM risk was analyzed by logistic regression and odds ratios are shown in table 5.1. Risk factors were identified based on the value of odds ratios. The higher risk was resulted from the association of gravidity (OR=1.5) and FHD (OR=1.84) with GDM. Consequently, these are the confounders that have been adjusted in the subsequent analyses.

Table 5.1: Assessment of the risk of GDM provided by the studied variables

BMI, Body mass index. DBP, diastolic blood pressure. FHD, Family history of diabetes.

Variables	OR	95% CI	*P* Value
Age, years	1.02	1.05-1.14	<0.001
BMI (kg/m^2)	1.04	1.05-1.16	<0.001
Gravidity	**1.5**	**1.05-2.17**	**0.03**
FHD	**1.84**	**1.28-2.66**	**<0.001**
DBP (mmHg)	1.02	1.00-1.04	0.03
Occupation (Others)	0.5	0.26-0.94	0.04

5.1.1 Association of the *CDKAL1* gene variants rs7756992 and rs7754840 with GDM

Crude analyses revealed that the altered allele G of rs7756992 significantly increase the odds of having GDM by more than 1.4 folds under codominant [AA vs. AG (OR = 1.56, 95% CI = 1.06 to 2.30), AA vs. GG (OR = 1.74, 95% CI = 0.88 to 3.45), ***P*=0.047**] dominant (OR = 1.59, 95% CI =1.10 to 2.30, ***P*= 0.014**) and log additive (OR=1.42, 95% CI=1.06 to 1.90, ***P*=0.019**) models. Though statistically insignificant the odds are also higher in both recessive (OR=1.37) and over dominant (OR=1.42) models (**Appendix A6: Table A1**). Adjustment for gravidity and FHD results significant associations of this variant with GDM under dominant (OR = **1.6**, 95% CI =1.07 to 2.29, *P*= **0.02**) and log additive (OR= **1.4,** 95% CI=1.05 to 1.91, *P*=**0.021**) models (**Figure 5.1**).

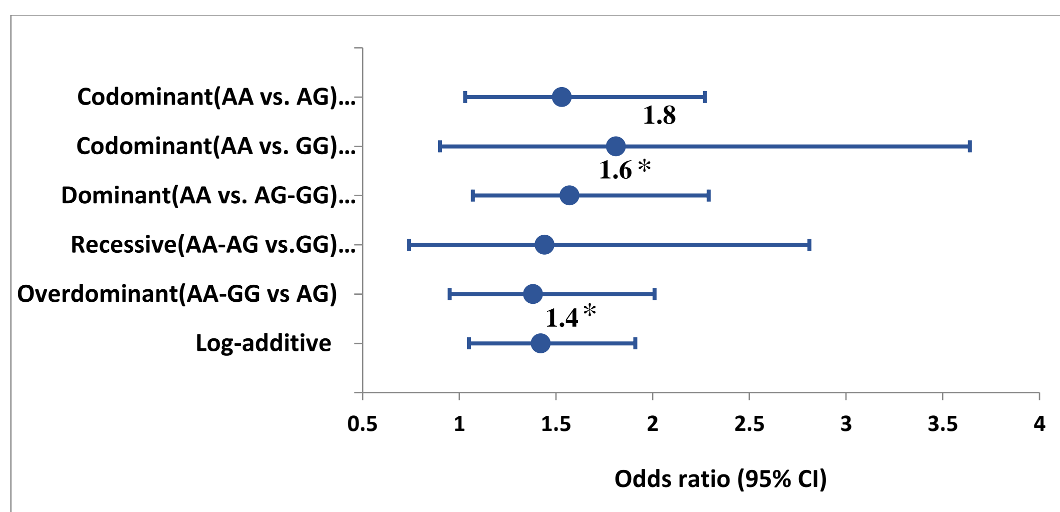


Figure 5.1: Associations of rs7756992(AA/AG/GG) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed circles and whiskers representing the 95% confidence intervals. (*, **, *** *P*<0.05, *P*<0.01, *P*<0.001)

On the other hand, no significant association was revealed by crude analyses between rs7754840 and GDM under any genetic model (**Appendix A6: Table A2**). When adjusted for confounding covariates, the odds of having GDM with CC genotype increased to 2.23 with a significance level of 95% CI from 1.04 to 4.75 under the codominant model. Impressively this adjustment results in significant (***P*=0.047**)

association between this SNP with GDM under the recessive model with an OR of 2.09 (95% CI= 1.00 to 4.36) (**Figure 5.2**).

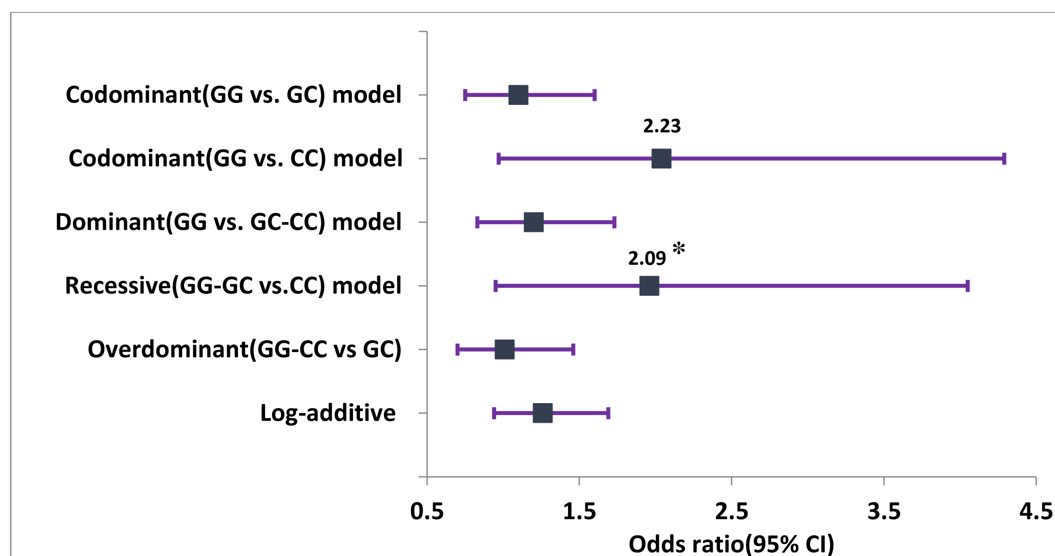


Figure 5.2: Associations of rs7754840(GG/GC/CC) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed squares and whiskers representing the 95% confidence intervals. (*, **, *** P<0.05, P<0.01, P<0.001)

5.1.2 Association of the *FTO* gene variant rs8050136 with GDM

Analyses revealed greater than 1.6 folds increase in odds of having GDM under codominant [C/C vs. A/A (OR = 1.70, 95% CI = 0.81 to 3.55), P=0.37] and recessive (OR = 1.63, 95% CI =0.80 to 3.33, P= 0.18) models after adjusting for FHD and gravidity. In the rest of the models, the odds remained between 1.1 and 1.2, with P values greater than 0.05 (**Figure-5.3**). None of the five genetic models showed a significant difference in the likelihood of developing GDM before this adjustment **(Appendix A6: Table A3).**

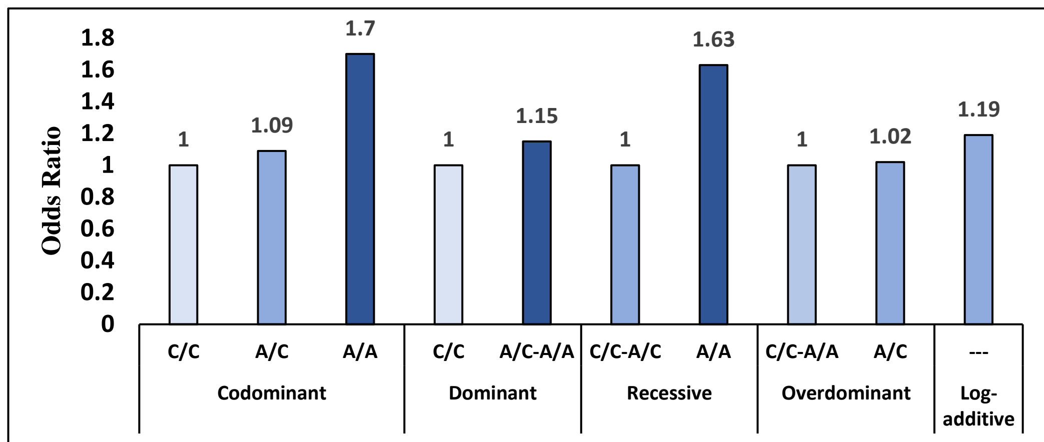


Figure 5.3: Associations of rs8050136 (CC/AC/AA) with GDM under different genetic models adjusted for family history of diabetes and gravidity.

5.1.3 Association of *HSPA1L* gene variant rs2227956 with GDM

There was no difference observed in odds ratios obtained from the three genotypes of this variant under any tested model before (**Appendix A6: Table A4**) and after adjustment (**Figure 5.4**) for confounders (FHD and gravidity). These observations were not statistically significant (P>0.05).

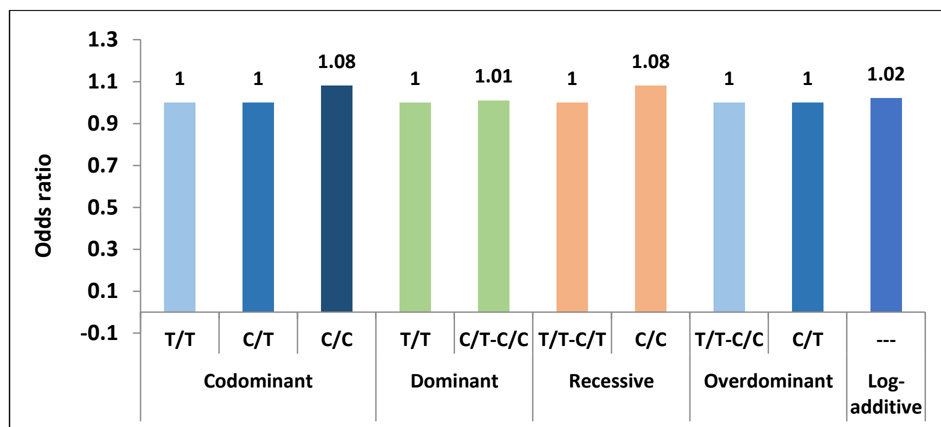


Figure 5.4: Associations of *HSPA1L* gene variant rs2227956 (TT/CT/CC) with GDM under different genetic models adjusted for family history of diabetes and gravidity

5.1.4 Association of *PPARG* gene rs3856806 and rs1801282 variants with GDM

Analyses of the rs3856806 association revealed odds of having GDM in both groups almost equal to 1 under all tested models **(Appendix A6: Table A5).** The odds

remained the same even after adjusting for the confounders (**Figure 5.5**) and these observations were statistically insignificant (*P*>0.05).

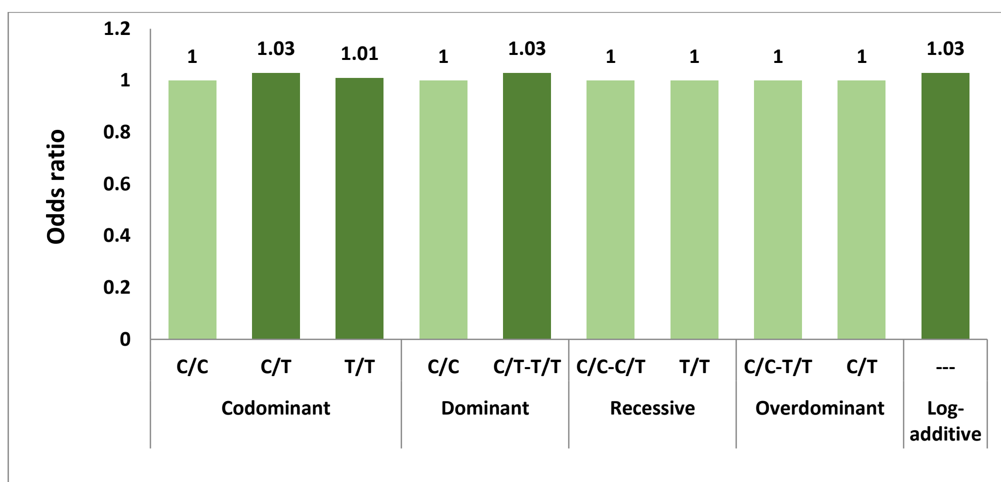


Figure 5.5: Associations of the *PPARG* gene rs3856806 (CC/CT/TT) with GDM under different genetic models adjusted for family history of diabetes and gravidity

Crude analyses revealed the GG genotype of altered allele G of rs1801282 decrease the odds of having GDM by 1.3 folds under codominant (OR = 0.75, 95% CI = 0.18-3.19) and recessive (OR = 0.78, 95% CI = 0.18-3.30) models (**Appendix A6: Table A6**). After adjustment for confounders, odds were decreased by 1.8 and 1.7 folds under codominant (OR = 0.57, 95% CI = 0.13-2.47) and dominant (OR = 0.58, 95% CI = 0.13-2.51) models respectively (**Figure 5.6**). The heterozygous genotype CG changed odds from 0.85 (95% CI = 0.55-1.30) to 0.90 (95% CI = 0.59-1.39) under dominant and from 0.85 (95% CI= 0.56-1.30) to 0.91(95% CI= 0.59-1.41) under overdominant models after adjustment for confounders (**Figure 5.6**). These decreases in odd ratios were not statistically significant (*P*>0.05).

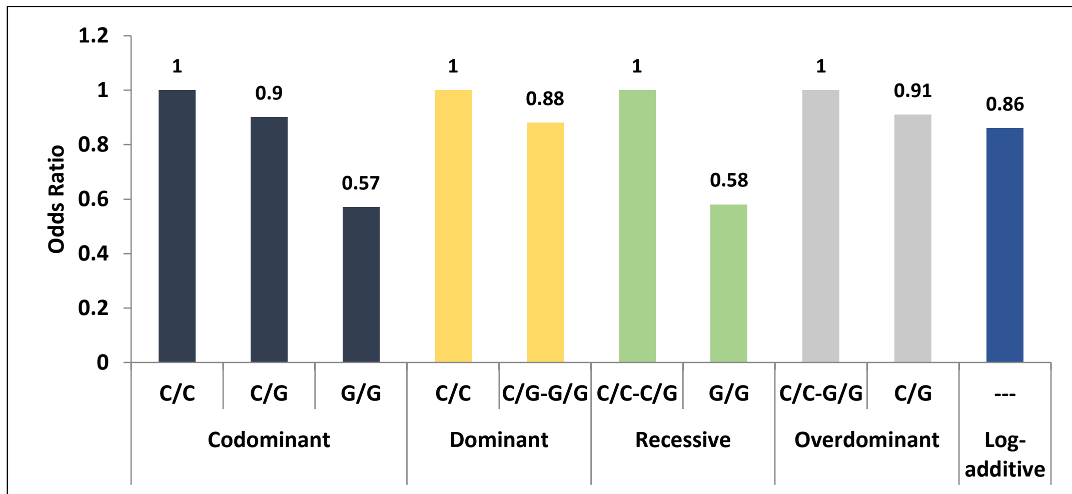


Figure 5.6: Associations of *PPARG* gene rs1801282 (CC/CG/GG) with GDM under different genetic models adjusted for family history of diabetes and gravidity

5.1.5 Association of the *TCF7L2* gene variant rs10885406 with GDM

Crude analyses revealed homozygous genotype of altered allele G decrease the odds of having GDM by 1.2 folds under codominant (OR = 0.86, 95% CI =0.47-1.56) and recessive (OR = 0.84, 95% CI= 0.48-1.47) models (**Appendix A6: Table A7**). After adjustment for confounders, odds were decreased by 1.3 folds under the same models (**Figure 5.7**). The heterozygous genotype AG has no impact on odds under dominant (OR = 1.01, 95% CI= 0.69-1.49) and overdominant (OR = 1.06, 95% CI= 0.74-1.53) models even after adjustment for confounders. These changes in odds under all tested models were not statistically significant (P>0.05)

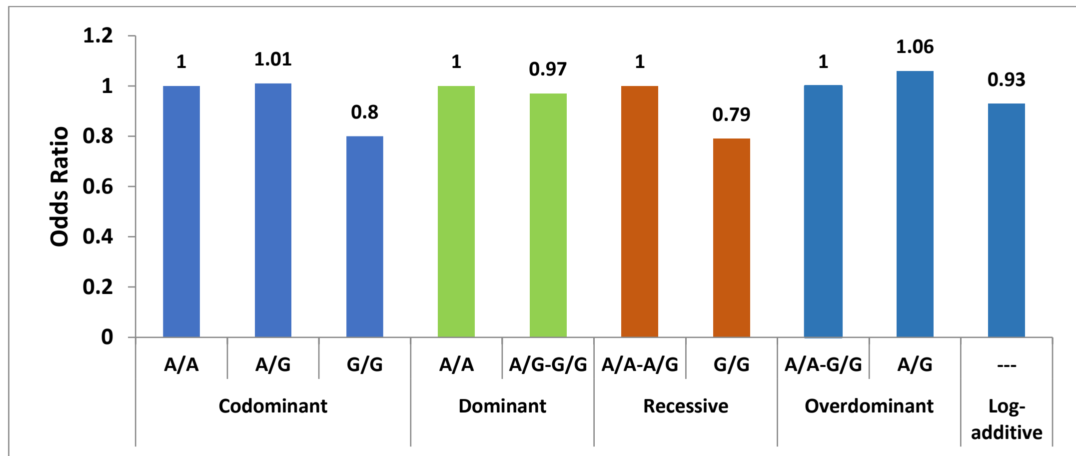


Figure 5.7: Associations of the *TCF7L2* gene rs10885406 (AA/AG/GG) with GDM under different genetic models adjusted for family history of diabetes and gravidity

5.1.6 Association of *TCF7L2* gene rs7903146 variant with GDM

The odds resulting from all three genotypes were the same under all tested models before and after adjustment (**Appendix A6: Table A8**). Analyses revealed altered allele T decreased the odds of having GDM by more than 1.4 and 1.3 folds under codominant (OR = 0.72, 95% CI = 0.33-1.56) and recessive (OR = 0.75, 95% CI = 0.35-1.58) models respectively (**Figure 5.8**). These decreases in odds were not statistically significant (*P*>0.05).

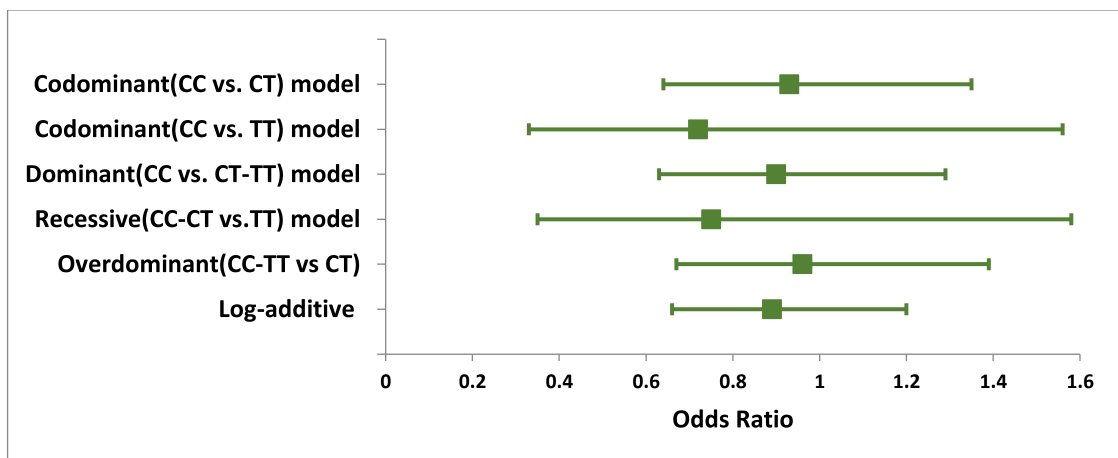


Figure 5.8: Associations of *TCF7L2* gene rs7903146 (CC/CT/TT) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed squares and whiskers representing the 95% confidence intervals.

5.1.7 Association of the *TCF7L2* gene variant rs12255372 with GDM

Analyses revealed altered allele T increased the odds of having GDM by more than 1.4 folds under all tested models, except under the recessive model (**Appendix A6: Table A9**). These increases were statistically significant under dominant (OR = 1.51, 95% CI = 1.06 -2.15, *P* = 0.023), overdominant (OR = 1.48, 95% CI = 1.04 to 2.11, *P* = 0.03) and log additive (OR = 1.37, 95% CI = 1.01-1.85, *P* = 0.041) models. After adjusting for the confounders, i.e., family history of diabetes and gravidity, the increase in odds by altered allele remained more or less the same, but these outcomes were statistically significant only under the dominant (OR = 1.44, 95% CI = 1.01- 2.07, *P* = 0.046) model (**Figure 5.9**).

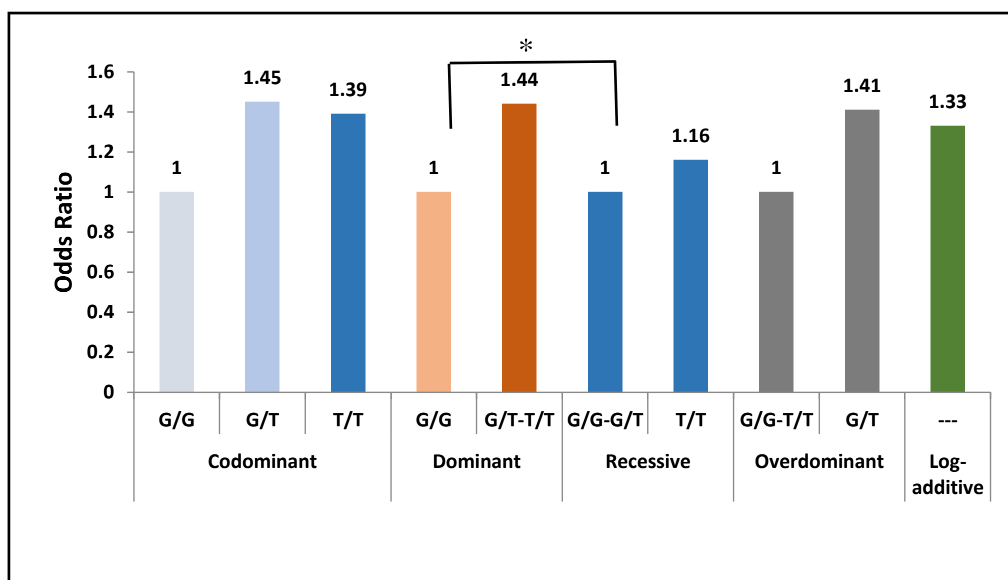


Figure 5.9: Associations of *TCF7L2* gene rs12255372 (GG/GT/TT) with GDM under different genetic models adjusted for family history of diabetes and gravidity. (*, **, *** *P*<0.05, *P*<0.01, *P*<0.001)

5.1.8 Association of the *WFS1* gene variant rs10010131 with GDM

Association analysis of the *WFS1* gene variant rs10010131 with GDM was carried out under all five genetic models and adjusted for family history of diabetes and gravidity (**Appendix A6: Table A10**). The odds increased very slightly from the same genotype in the case of AA (1.08 times), but no changes resulted in the GG and AG genotypes

after adjustment (**Figure 5.10**). Under codominant and recessive models, the AA genotype increases the odds of GDM by 1.5 (95% CI = 0.74 to 3.04), and 1.61 (95% CI = 0.81 to 3.22) folds from the GG genotype, respectively, after adjusting for confounding factors. The heterozygous genotype of this variant decreases the chances of having GDM when analyzed under codominant, dominant, and overdominant models.

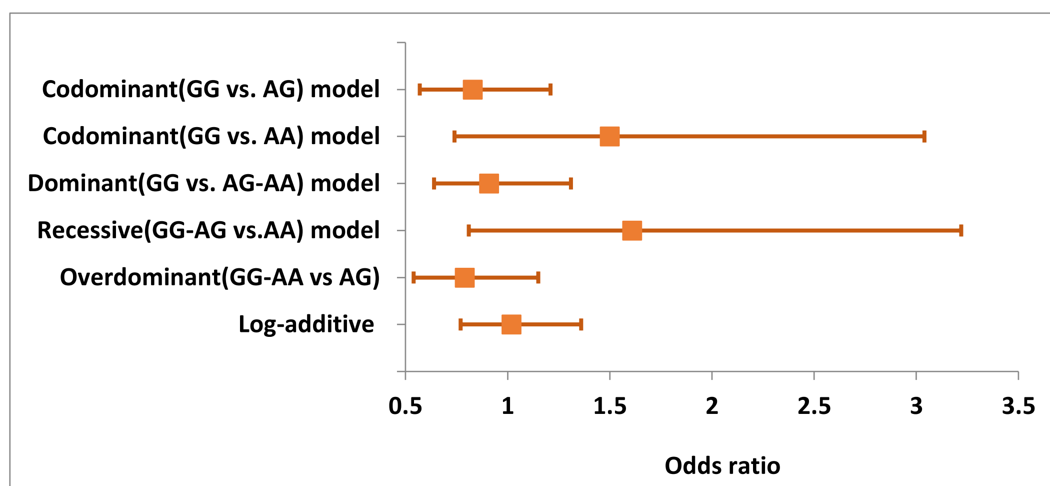


Figure 5.10: Associations of rs10010131 (GG/AG/AA) with GDM under different genetic models adjusted for family history of diabetes and gravidity with odds ratios shown by closed squares and whiskers representing the 95% confidence intervals.

The observed changes in odds of having GDM under all tested models were not statistically significant (*P*>0.05).

5.2 Association of the alleles of the selected T2DM linked genetic variants with GDM

Four of the 10 examined T2DM linked variants showed an odds ratio (OR) greater than 1 for the GDM group compared with the control group ranging from 1.2 to 1.36 (**Table 5.2**). In presence of the G allele of rs7756992 of the *CDKAL1* gene, the odds of having GDM increase significantly (***P*=0.03**) by 1.36-fold (95% CI=1.031 to 1.782). Though the frequency of the C allele of rs7754840 was higher in the GDM group (**Table 4.3**) the effect of this allele on the susceptibility of GDM is insignificant (OR=1.25, 95% CI=0.939 to 1.675, *P*=0.12). In the presence of the risk alleles of the *FTO* and the *TCF7L2* gene variant, the odds of having GDM increase by 1.2 and 1.31 folds respectively. The allele frequencies for three of the target variations were not associated

[rs2227956 (*HSPA1L*): OR = 1.01, rs3856806 (*PPARG*): OR = 1.00, rs1001013 (*WFS1*): OR= 1.00] with a risk for GDM. The odds of having GDM decreased by the altered allele of the three of the studied SNPs [rs1801282 (*PPARG*): OR=0.85, rs7903146 (*TCF7L2*): OR=0.93, rs10885406 (*TCF7L2*): OR=0.96] (**Table 5.2**).

Table 5.2: Association of the allele frequencies of the target SNPs with GDM

Gene	SNP	Allele	Control	GDM	OR	95% CI	*P* Value
CDKAL1	rs7756992	A	360	269	1.36	1.031 to 1.782	**0.03**
		G	153	155			
	rs7754840	G	384	299	1.25	0.939 to 1.675	0.12
		C	128	125			
FTO	rs8050136	C	412	300	1.20	0.9104 to 1.5744	0.20
		A	156	136			
HSPA1L	rs2227956	T	486	374	1.01	0.7040 to 1.4406	0.97
		C	80	62			
PPARG	rs1801282	C	494	386	0.85	0.5862 to 1.2418	0.41
		G	78	52			
	rs3856806	C	469	357	1.00	0.7223 to 1.3888	0.99
		T	101	77			
TCF7L2	rs7903146	C	406	317	0.93	0.6994 to 1.2233	0.58
		T	162	117			
	rs12255372	G	426	301	1.31	0.9898 to 1.7263	**0.059**
		T	144	133			
	rs10885406	A	362	281	0.96	0.7434 to 1.2478	0.78
		G	210	157			
WFS1	rs1001013	G	419	324	1.00	0.7547 to 1.3328	0.98
		A	147	114			

5.3 Association haplotypes of T2DM linked genetic variants with GDM

Haplotype analysis provided further evidence of association by showing a significant difference between cases and controls (1). The target SNPs located in the same chromosome were analyzed for the association of their haplotypes with GDM.

5.3.1 The Association of Haplotype of *CDKAL1* gene variants, rs7756992 and rs7754840, with GDM

Multivariate analysis adjusted for gravidity and family history of diabetes validated the relationship of the rs7754840 C-and rs7756992 G-allele containing (CG) haplotype with GDM, conferring significant (P=0.032) disease susceptibility with an odd of 1.43. (1.03-1.98) (**Table 5.3**).

Table 5.3: The association of haplotype of rs7756992 and rs7754840 with GDM

Sl.	rs7754840	rs7756992	Frequency	OR (95% CI)	*P*-value [a]
1	**G**	**A**	0.6417	1.00	
2	**C**	**G**	0.243	**1.43(1.03-1.98)**	**0.032**
3	**G**	**G**	0.0857	1.40 (0.84 - 2.33)	0.19
4	**C**	**A**	0.0296	1.15 (0.50 - 2.67)	0.74

[a] adjusted for gravidity and family history of diabetes

5.3.2 Association analyses of Haplotypes of *CDKAL1* and *HSPAL1* gene variants with GDM

The *CDKAL1* and *HSPAL1* genes were located on the same chromosome (**Table 3.1**). For that reason, association of the haplotypes of the target SNPs of these genes were analyzed and the results obtained were shown in the **Table 5.4**. The common haplotype AGT was used as reference (OR = 1.00). The risk alleles containing haplotype GCC increases the odds of having GDM by 2.8 folds.

Table 5.4 The association of Haplotypes of *CDKAL1* and *HSPAL1* gene variants with GDM(n=466)

Sl.	rs7754840	rs7756992	rs2227956	Frequency	OR (95% CI)	*P*-value [a]
1	A	G	T	0.5408	1.00	---
2	G	C	T	0.2116	1.27 (0.87 - 1.86)	0.21
3	A	G	C	0.1033	0.88 (0.53 - 1.47)	0.64
4	G	G	T	0.0742	1.38 (0.78 - 2.44)	0.27
5	A	C	T	0.0297	1.08 (0.46 - 2.51)	0.87
6	**G**	**C**	**C**	**0.027**	**2.77 (0.97 - 7.90)**	**0.057**
7	G	G	C	0.0134	1.47 (0.35 - 6.14)	0.6

[a] adjusted for gravidity and family history of diabetes

5.3.3 The Association of the Haplotypes of the *PPARG* gene variants rs3856806 and rs1801282 with GDM

Using the common rs3856806/ rs1801282 CC haplotype as reference (OR = 1.00), multivariate analysis adjusted for gravidity and family history of diabetes revealed protection against GDM by rs3856806 T- and rs1801282 G-allele containing (TG) haplotype with an odd of 0.85 (95% CI= 0.55 - 1.31) (**Table 5.5**).

Table 5.5: The association of haplotype of rs3856806 and rs1801282 with GDM

Sl.	rs3856806	rs1801282	Frequency	OR (95% CI)	*P*-value [a]
1	C	C	0.798	1.00	---
2	**T**	**G**	**0.1037**	**0.85 (0.55 - 1.31)**	**0.46**
3	T	C	0.0736	1.29 (0.80 - 2.08)	0.29
4	C	G	0.0248	1.11 (0.49 - 2.52)	0.8

[a] adjusted for gravidity and family history of diabetes

5.3.4 The Association of Haplotype of *TCF7L2* gene variants with GDM

For haplotype analysis of the target SNPs of *TCF7L2* gene, adjusted multivariate logistic regression was carried out by using the GAC haplotype as reference (OR=1.00). The altered allele containing haplotype TGT insignificantly increase the odds of GDM by 1.2 folds.

Table 5.6: The association of haplotype of *TCF7L2* gene variants with GDM

Sl.	rs12255372	rs10885406	rs7903146	Frequency	OR (95% CI)	*P*-value [a]
1	G	A	C	0.52	1.00	---
2	G	G	T	0.1415	0.87 (0.55 - 1.38)	0.57
3	**T**	**G**	**T**	**0.1304**	**1.15 (0.77 - 1.72)**	**0.5**
4	T	A	C	0.1076	1.55 (0.89 - 2.69)	0.13
5	G	G	C	0.06	0.97 (0.50 - 1.86)	0.92
6	T	G	C	0.0328	1.23 (0.51 - 2.95)	0.64
Rare	*	*	*	0.0077	0.25 (0.03 - 2.14)	0.21

[a] adjusted for gravidity and family history of diabetes

5.4 Summary

- The association analyses of the target SNPs with GDM were carried out under **five genetic models**; codominant, dominant, recessive, overdominant and log-additive models.
- Family history of diabetes and gravidity were detected as **confounder variables** and have been adjusted in the subsequent analyses.
- The *CDKAL1* gene variants, **rs7756992** and **rs7754840,** and the *TCF7L2* gene variant **rs12255372** were **significantly** associated with the susceptibility of GDM.
- The nature of the association of the target SNPs are shown in **Table 5.7**. The odds of having GDM from best fitted (lowest AIC and BIC values) models (Section 3.12) are noted here.

Table 5.7: Nature of association of the target SNPs with GDM.

Gene	SNP ID	Odds ratio	Nature of Association
CDKAL1	rs7756992	1.56	Susceptible
	rs7754840	2.09	Susceptible
FTO	rs8050136	1.63	Susceptible
HSPA1L	rs2227956	1.01	No association
PPARG	rs1801282	0.58	Protective
	rs3856806	1.03	No association
TCF7L2	rs10885406	0.83	Protective
	rs7903146	0.77	Protective
	rs12255372	1.44	Susceptible
WFS1	rs1001013	1.61	Susceptible

OR=1(or close to 1.0), the SNP is not associated with the disease; OR is greater than 1.0, the SNP is might be a risk factor; OR is less than 1.0, the SNP might be a protective factor

- Only the risk allele of the *CDKAL1* gene variants, rs7756992, **significantly increased** the chances of GDM.
- The risk alleles containing (CG) haplotype of the *CDKAL1* gene variants, rs7756992 and rs7754840 conferred **significant (*P*=0.032) disease susceptibility** with an odd of 1.43(1.03-1.98).

- The risk alleles containing haplotype (GCC) of the *CDKAL1* and *HSPA1L* gene variants, rs7754840, rs7756992 and rs2227956, increased the susceptibility of GDM by 2.8 folds.

Reference

1. Carter AM, Standeven KF, Grant PJ. Chapter 54 - Common Genetic Determinants of Coagulation and Fibrinolysis. In: Rimoin D, Pyeritz R, Korf B, editors. Emery and Rimoin's Principles and Practice of Medical Genetics (Sixth Edition). Oxford: Academic Press; 2013. p. 1-20.

6. Cumulative association of family history of diabetes and selected SNPs with the predisposition of GDM

6.1 Association of target SNPs and family history of Diabetes with GDM

Selected SNPs have been reported for their association with T2DM in several studies in different populations (1-10). Moreover, family history of diabetes is a confounding variable in this study. Confounding can be managed, unlike other types of biases, by controlling for it after a study is completed using stratification (11), so the individuals in this study were separated into two strata: positive family history of diabetes and no family history of diabetes (12). Relative risk analysis for each stratum was carried out for control and GDM groups by cross classification interactions using multivariate logistic regression under four genetic (codominant, dominant, recessive, and overdominant) models adjusted for gravidity. The cumulative impact of these variants and PFHD on GDM has been detected.

6.1.1 Association of the *CDKAL1* Gene Variants with the Family History of Diabetes

The cross-classification interaction of AA genotype of the rs7756992 results 1.2 folds increase in odds of having GDM with a PFHD under codominant and dominant models (**Figure 6.1**). The heterozygous (AG) and homozygous (GG) genotypes of the risk allele G increase the odds by 2.5(OR=2.93) and 3.7 (OR=4.07) folds under the codominant model, respectively. Under the recessive model, the GG genotype increases the odds by 3.8 folds (OR=4.11) (**Appendix A7: table A11**).

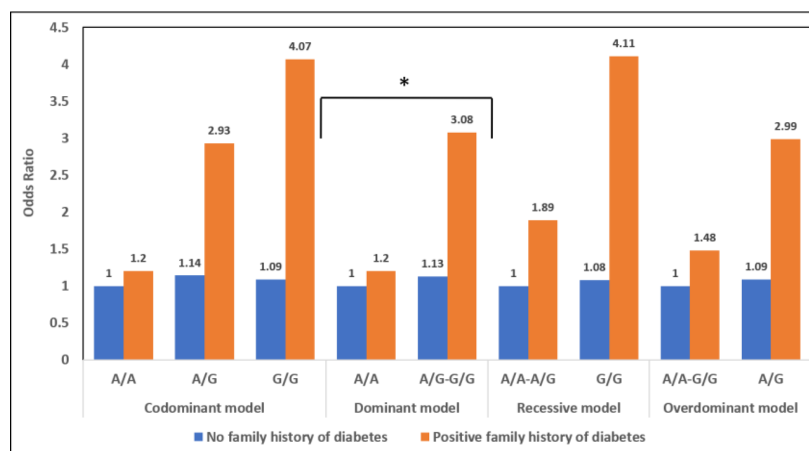


Figure 6.1: Correlation between GDM Risk with cumulation of the *CDKAL1* gene variant rs7756992 polymorphism, and family history of diabetes. (*, **, *** *P*<0.05, *P*<0.01, *P*<0.001)

Among the tested models, significant (P=0.038) increase in the odds of GDM observed only under the dominant model.

In the presence of PFHD, the CC genotype of rs7754840 increase the odds of GDM by 4.4 and 4.8 folds under codominant and recessive models, respectively. The reference genotype GG increases the odds by 1.7 folds under codominant and dominant models. The heterozygous genotype GC of this variant also increases the odds by 2.2 folds under codominant and overdominant models (**Figure 6.2**).

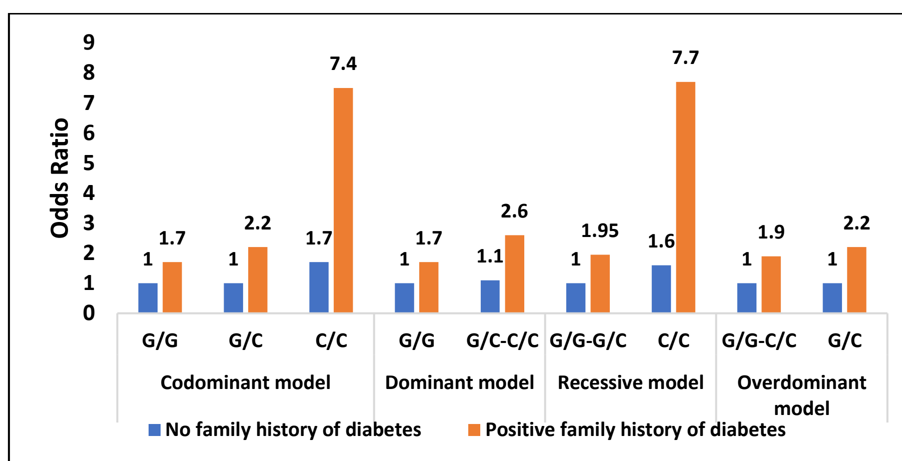


Figure 6.2: Correlation between GDM Risk with cumulation of the *CDKAL1* rs7754840 polymorphism, and family history of diabetes under four genetic models adjusted for gravidity.

None of these changes in odds of GDM were statistically significant (P>0.05).

6.1.2 Interaction between the *FTO* gene variant rs8050136 and family history of diabetes on the susceptibility of GDM

The combined impact of family history of diabetes and the rs8050136 variant of the *FTO* gene resulted in increased odds of having GDM under all tested genetic models adjusted for the gravidity. Even in individuals with a CC genotype, the odds increased by 2.22 folds (95%CI = 1.3–3.8) under codominant and dominant models (**Figure 6.3**). In the presence of PFHD and the AA genotype, the risk of GDM increased from 2.26 to 2.67 and from 2.04 to 2.40) under codominant and recessive models, respectively (**Appendix A7: Table A12**).

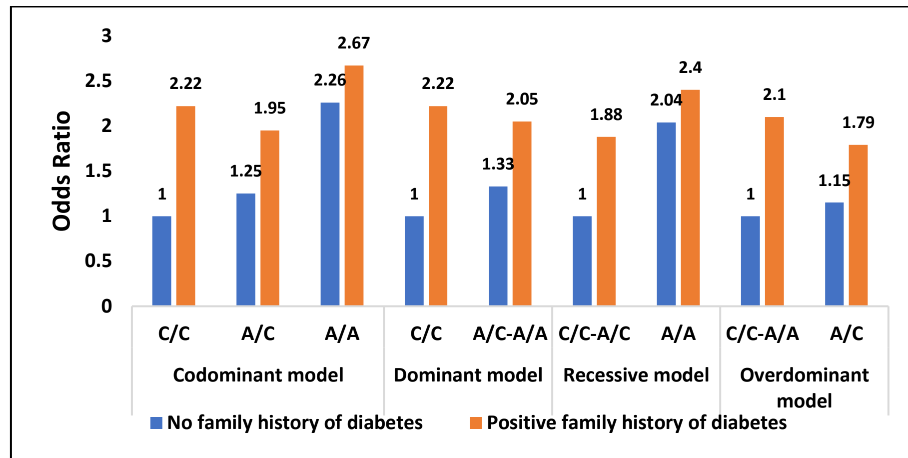


Figure 6.3: Correlation between GDM Risk with cumulation of the *FTO* gene rs8050136 polymorphism, and family history of diabetes under different genetic models and adjusted for gravidity.

Though the odds of GDM increased substantially, these were not statistically significant ($P > 0.05$) under any tested model.

6.1.3 Interaction between the *HSPA1L* gene rs2227956 and family history of diabetes on the susceptibility of GDM

The cumulative effect of family history of diabetes and wild type allele T of the rs2227956 variant of the *HSPA1L* gene resulted in increased odds of having GDM under all tested genetic models adjusted for gravidity (**Figure 6.4**). For individuals with the TT genotype, the odds increased by more than 2 folds under codominant and dominant models. In the presence of PFHD and the CC genotype of this variant, the risk of GDM decreased from 2.38 to 0.91 and from 2.27 to 0.87 under codominant and recessive models, respectively (**Figure 6.4**).

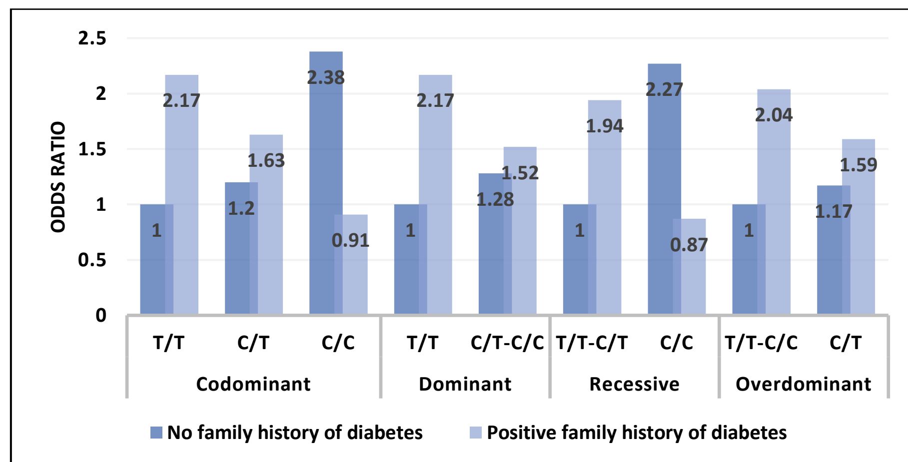


Figure 6.4: Association between GDM Risk with cumulation of the *HSPAL1* gene rs2227956 polymorphism, and family history of diabetes.

Though the odds of GDM changed substantially, they were not statistically significant ($P > 0.05$) under any tested model (**Appendix A7: Table A13**).

6.1.4 Interaction between the *PPARG* gene variants (rs3856806 and rs1801282) and family history of diabetes on the susceptibility of GDM

A positive family history of diabetes increased the odds of GDM under all tested genetic models adjusted for gravidity. Even in individuals with the CC genotype of the reference allele C of rs3856806, there were 2 folds higher odds under codominant and dominant models. In the presence of PFHD and the TT genotype of this variant, the risk of GDM increased from 1.11 to 1.82 and from 1.07 to 1.76 under codominant and recessive models, respectively (**Appendix A7: Table A14**). The heterozygous CT genotype also increased the odds from 1.12 to 1.8 under codominant and overdominant models (**Figure 6.5**).

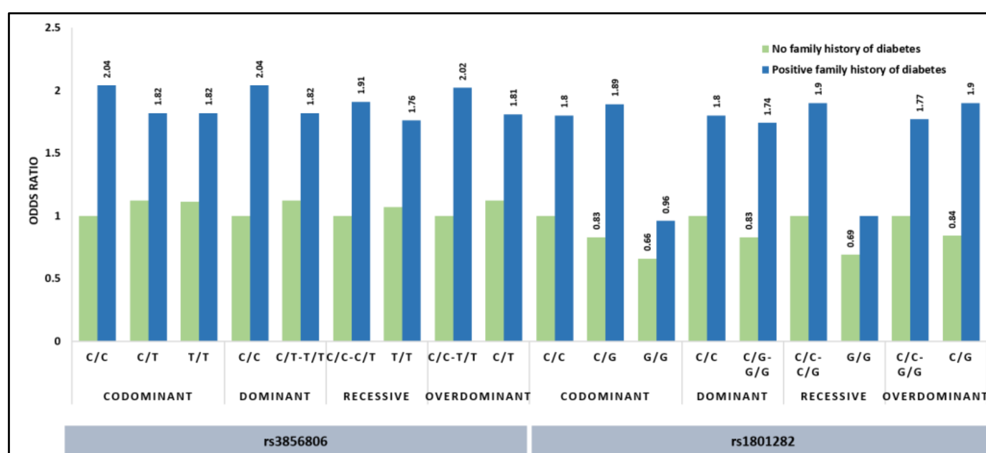


Figure 6.5: Cumulative association of GDM Risk of the *PPARG* gene polymorphisms, and family history of diabetes.

Individuals having PFHD with CC genotype of the reference allele C of rs1801282 resulted in 1.8 folds higher odds under codominant and dominant models (**Figure 6.6**). In the presence of PFHD and the GG genotype of this variant, the risk of GDM increased from 0.66 to 0.96 and from 0.69 to 1.00 under codominant and recessive models, respectively. The heterozygous CG genotype also increased the odds from 0.83 to 1.89 and from 0.84 to 1.90 under codominant and overdominant models. Though the odds of GDM changed substantially in the case of both variants, these were not statistically significant ($P > 0.05$) under any tested model (**Appendix A7: Table A15**).

6.1.5 Association of the *TCF7L2* gene variant rs10885406 with the Family History of Diabetes

Homozygous genotype AA of reference allele A significantly increased the chnaces of having GDM in the presence of PFHD by 3.4 folds under codominant and dominant models (**Figure 6.6**). The heterozygous AG genotype also resulted in significantly higher odds in this group containing PFHD under codominant (OR = 1.95) and overdominant (OR = 1.88) models (**Appendix A7: Table A16**). The presence of PFHD and GG genotype increased the odds from 1.20 to 1.63 and from 0.97 to 1.31 under the codominant and recessive models, respectively (**Figure 6.6**).

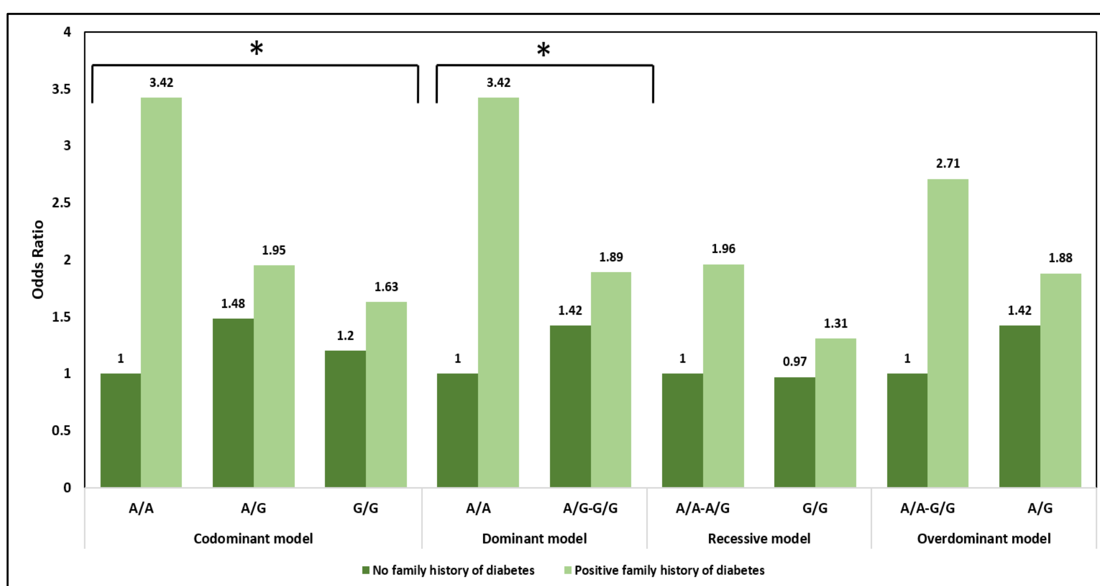


Figure 6.6: Cumulative Correlation between GDM Risk of the *TCF7L2* gene rs10885406 polymorphism, and family history of diabetes. (*, **, *** $P<0.05$, $P<0.01$, $P<0.001$)

6.1.6 Association of the *TCF7L2* gene rs7903146 variant with the Family History of Diabetes

The presence of PFHD increased the odds of having GDM under all tested genetic models adjusted for gravidity (**Figure 6.7**). Even in individuals with CC genotype of reference allele C resulted in 2 folds higher odds under codominant and dominant models. In the presence of PFDH and the TT genotype, the risk of GDM increased from 0.71 to 1.46 under codominant and recessive models, respectively (**Appendix A7: Table A17**). The heterozygous CT genotype also increased the odds from 1.01 to 1.69 and from 1.04 to 1.74 under codominant and overdominant models, respectively.

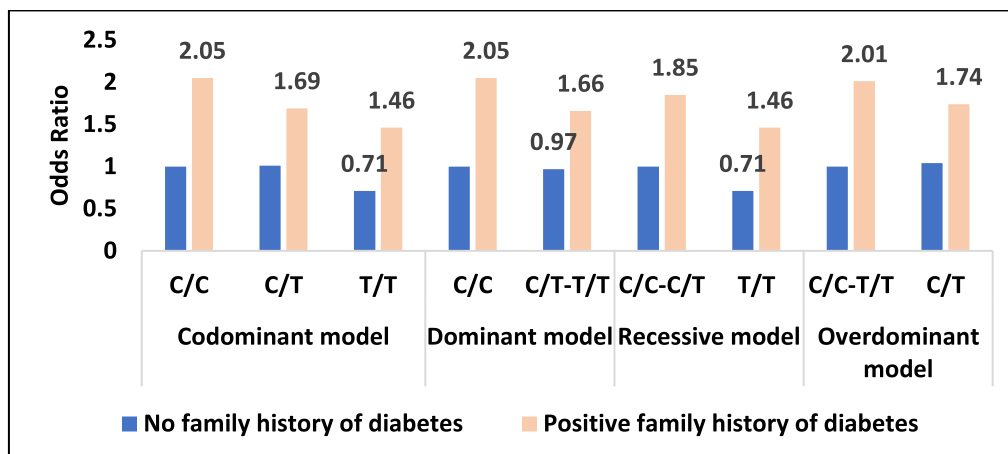


Figure 6.7: Cumulative correlation between GDM Risk of the *TCF7L2* gene rs7903146 polymorphism, and family history of diabetes.

Though the odds of GDM changed substantially, these were not statistically significant(P>0.05) under any tested model.

6.1.7 Association of the *TCF7L2* gene rs12255372 variant with the Family History of Diabetes

The presence of PFHD increased the odds of having GDM under all tested genetic models adjusted for gravidity (**Figure 6.8**). Even in individuals with the GG genotype of reference allele G, the odds increased by 1.7 folds under codominant and dominant models. In the presence of PFHD and the TT genotype of this variant, the risk of GDM increased from 1.05 to 3.8 and from 0.88 to 3.21 under codominant and recessive models, respectively. The heterozygous GT genotype also increased the odds from 1.45 to 2.49 and from 1.45 to 2.48 under codominant and overdominant models.

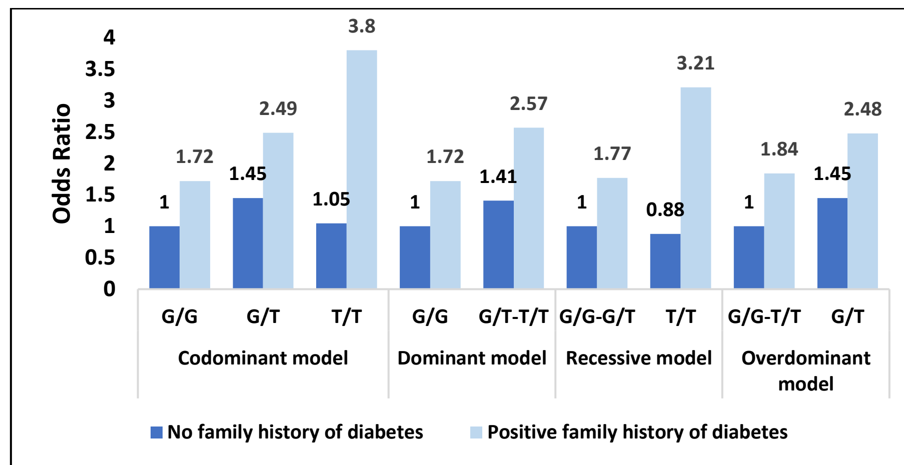


Figure 6.8: Cumulative association of GDM Risk with the *TCF7L2* gene rs12255372 polymorphism, and family history of diabetes.

Though the odds of GDM changed substantially, these were not statistically significant(P>0.05) under any tested model (Appendix A7: Table A18).

6.1.8 Association of the *WFS1* gene variant rs10010131 with the Family History of Diabetes

In individuals with GG genotype of the reference allele G of rs10010131, the odds increased by 2.3 folds under codominant and dominant models (**Figure 6.9**). In the presence of the PFHD and the AA genotype of this variant, the risk of GDM decreased from 1.98 to 1.79 and from 2.01 to 1.81 under codominant and recessive models, respectively (**Appendix A7: Table A19**). The heterozygous AG genotype also increased the odds from 0.96 to 1.53 and from 0.87 to 1.40 under codominant and overdominant models, respectively.

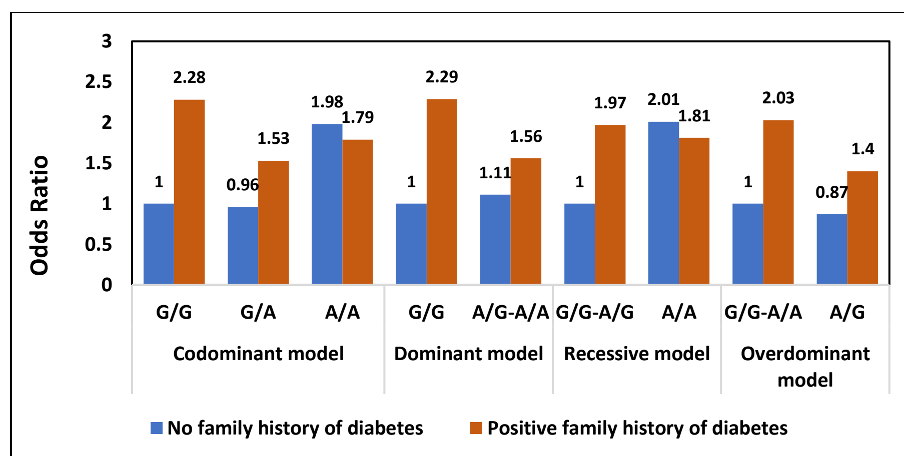


Figure 6.9: Cumulative association of GDM Risk with the *WFS1* gene rs10010131 polymorphism, and family history of diabetes.

Though the odds of GDM changed substantially, these were not statistically significant (*P*>0.05) under any tested model.

6.2 Summary

- The cumulative impact of target SNPs and family history of diabetes on GDM were analyzed under **four genetic models**; codominant, dominant, recessive and overdominant models.
- These interaction analyses were adjusted for gravidity.
- **Significant susceptible association** with GDM in presence of positive family history of diabetes was revealed from the interaction analyses of the ***CDKAL1* gene variant rs7756992** and **the *TCF7L2* gene variant rs10885406**.
- **Protective association** (2.6 folds decrease in odds of having GDM) in presence of positive family history of diabetes was revealed by the risk allele C of the *HSPA1L* gene variant rs2227956 under codominant and recessive models. But this association was not statistically significant.
- In case of rest of the SNPs, women with family history of diabetes had about 1.5 to 4 folds higher odds of GDM. The results did not reach statistical significance.

Reference

1. Bakhashab S, Filimban N, Altall RM, Nassir R, Qusti SY, Alqahtani MH, et al. The Effect Sizes of PPARγ rs1801282, FTO rs9939609, and MC4R rs2229616 Variants on Type 2 Diabetes Mellitus Risk among the Western Saudi Population: A Cross-Sectional Prospective Study. Genes (Basel). 2020;11(1):98.
2. El-Lebedy D, Ashmawy IJJoGE, Biotechnology. Common variants in TCF7L2 and CDKAL1 genes and risk of type 2 diabetes mellitus in Egyptians. 2016;14(2):247-51.
3. Han X, Luo Y, Ren Q, Zhang X, Wang F, Sun X, et al. Implication of genetic variants near SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, FTO, TCF2, KCNQ1, and WFS1 in type 2 diabetes in a Chinese population. 2010;11(1):1-9.

4. Scott LJ, Bonnycastle LL, Willer CJ, Sprau AG, Jackson AU, Narisu N, et al. Association of Transcription Factor 7-Like 2 (TCF7L2) Variants With Type 2 Diabetes in a Finnish Sample. Diabetes. 2006;55(9):2649-53.
5. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007;445(7130):881-5.
6. Xiao S, Zeng X, Quan L, Zhu J. Correlation between polymorphism of FTO gene and type 2 diabetes mellitus in Uygur people from northwest China. Int J Clin Exp Med. 2015;8(6):9744-50.
7. Li T, Xu X, Xu Y, Jin P, Chen J, Shi Y, et al. PPARG Polymorphisms Are Associated with Unexplained Mild Vision Loss in Patients with Type 2 Diabetes Mellitus. Journal of ophthalmology. 2019;2019:5284867.
8. Moniruzzaman M, Ahmed I, Huq S, All Mahmud MS, Begum S, Mahzabin Amin U, et al. Association of polymorphism in heat shock protein 70 genes with type 2 diabetes in Bangladeshi population. 2020;8(2):e1073.
9. Wang J, Kuusisto J, Vänttinen M, Kuulasmaa T, Lindström J, Tuomilehto J, et al. Variants of transcription factor 7-like 2 (TCF7L2) gene predict conversion to type 2 diabetes in the Finnish Diabetes Prevention Study and are associated with impaired glucose regulation and impaired insulin secretion. Diabetologia. 2007;50(6):1192-200.
10. Lv X, Zhang L, Sun J, Cai Z, Gu Q, Zhang R, et al. Interaction between peroxisome proliferator-activated receptor gamma polymorphism and obesity on type 2 diabetes in a Chinese Han population. Diabetol Metab Syndr. 2017;9:7.
11. Jager K, Zoccali C, Macleod A, Dekker FJKi. Confounding: what it is and how to deal with it. 2008;73(3):256-60.
12. Amin USM, Parvez N, Rahman TA, Hasan MR, Das KC, Jahan S, et al. CDKAL1 gene rs7756992 A/G and rs7754840 G/C polymorphisms are associated with gestational diabetes mellitus in a sample of Bangladeshi population: implication for future T2DM prophylaxis. Diabetol Metab Syndr. 2022;14(1):18.

7. Cumulative association of target SNPs and Gravidity with GDM

7.1 Cumulative association of target SNPs and Gravidity with GDM

GDM showed an association with high gravida in various studies (1-4). Since gravidity has a confounding effect on the association with GDM in this study, participants were divided into two strata; multigravida and primigravida, to analyze the cumulative impact of this variable and target SNPs on GDM. Relative risk analysis for each stratum was conducted for control and case groups by cross classification interactions using multivariate logistic regression under four genetic (codominant, dominant, recessive, and overdominant) models adjusted for family history of diabetes. The cumulative impact of the target SNPs and gravidity on GDM has been detected.

7.1.1 Association of the *CDKAL1* Gene Variants with Gravidity

Interaction analyses for the variants of the *CDKAL1* gene revealed an increase in the odds of GDM in multigravida women. The AA genotype of rs7756992 increased odd ratios by 1.4 (OR=1.35) folds in multigravida women under codominant and dominant models, whereas for GG genotype risk of having GDM increased by more than 2 folds (OR= 2.09 and OR=2.11) in both strata [Figure 7.1(a) and Appendix A8: Table A20]. There was a distinct increase observed in the odds of having GDM compared to primigravida in the case of rs7754840, with odds varied from 1.39 (95% CI= 0.83-2.32) to 3.05 (95% CI= 1.12-8.30) under the codominant and dominant models [Figure 7.1(b) and Appendix A8: Table A20].

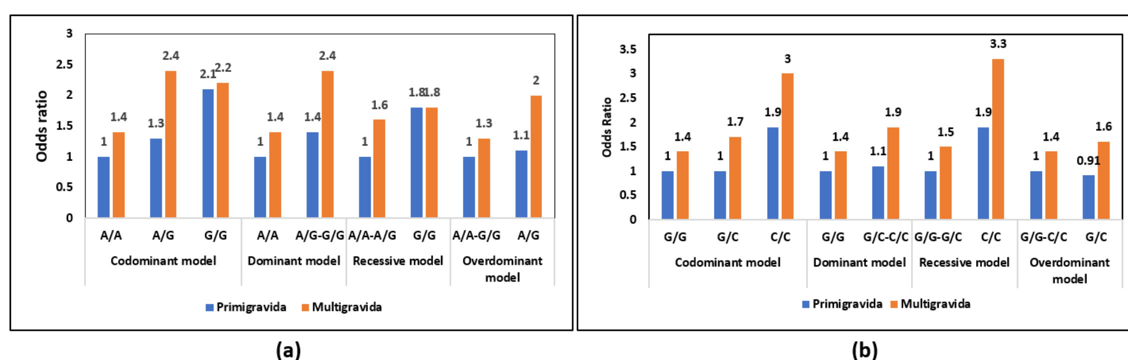


Figure 7.1: Association of GDM Risk with cumulation of the *CDKAL1* gene variants (a) rs7756992 (b) rs7754840, and gravidity.

7.1.2 Multigravidity and the *FTO* gene variant rs8050136 cumulatively increase chances of GDM

Interaction analyses of the *FTO* gene variant rs8050136, gravidity, and GDM revealed significant changes in odd ratios under all genetic models adjusted for family history of diabetes except under the recessive model. Codominant (OR=0.53,95%CI=0.29-0.95, *P*=0.0068) and overdominant (OR=0.52, 95%CI=0.29-0.92, *P*=0.0025) models revealed significant protective association of AC genotype with GDM in primigravida group. This protective role was also observed under the dominant model (OR=0.59, 95%CI=0.33-1.03) in the same group and was statistically significant(*P*=0.0021). The CC genotype of reference allele showed significant protection against GDM in multigravida under codominant (OR=0.84, 95%CI=0.50-1.42, *P*=0.0068), dominant (OR=0.84, 95%CI=0.50-1.42, *P*=0.0068), and overdominant (OR=0.92, 95%CI=0.57-1.49, *P*=0.0025) models (**Appendix A8: Table A21**). The risk of having GDM increased significantly by 1.6 to 2.3 folds under these three models in the presence of this variant in multigravida group. The chances of GDM in multigravid women increased significantly (*P*=0.0068) by 1.8(OR=1.51) folds in the presence of AC and 2.3(OR=1.96) folds in the presence of AA genotypes. Compared to primigravid women, the AC genotype significantly increases the odds of GDM in multigravid women by 2.8 and 2.9 folds under codominant and overdominant models, respectively. On the other hand, the AA genotype increases odds by 1.8 folds under codominant and recessive models (**Appendix A8: Table A21**). Analysis by recessive model showed highest odds (OR=2.64) though statistically insignificant.

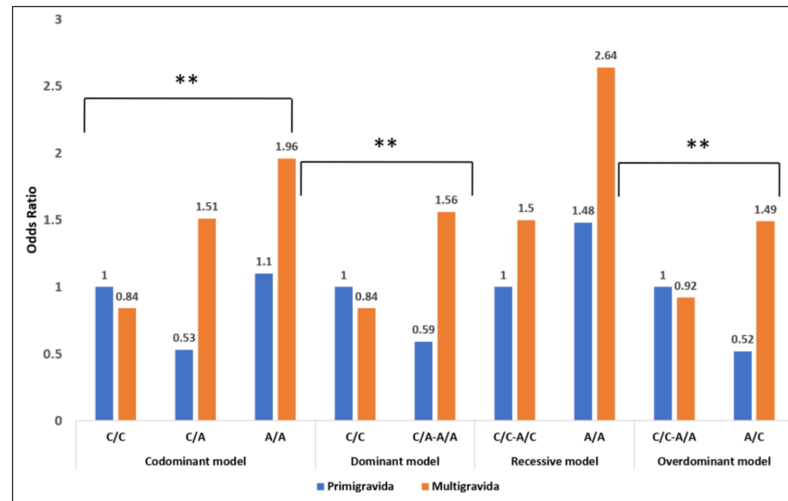


Figure 7.2: Correlation between GDM Risk with cumulation of the *FTO* gene rs8050136 polymorphism, and gravidity. (*, **, *** $P<0.05$, $P<0.01$, $P<0.001$)

7.1.3 Multigravidity and the *HSPA1L* gene variant rs2227956 cumulatively increase chances of GDM

Within the primigravida group, the CC genotype increased the chances of having GDM by 1.7 folds under the codominant and recessive models and decreased the odds of GDM 1.28 times in multigravida from primigravida group under same models. The TT genotype in multigravida women has 1.59 folds higher odds of GDM than primigravida women under the codominant and dominant models (**Appendix A8: Table A22**).

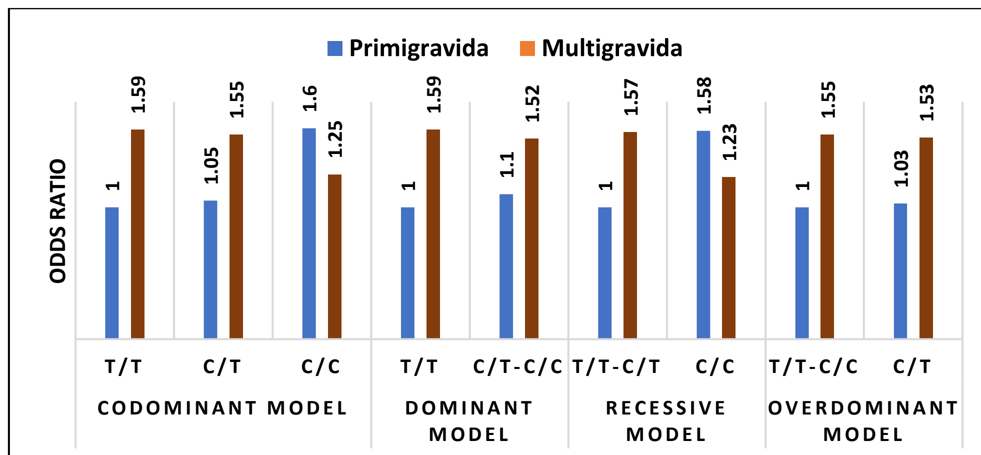


Figure 7.3: Correlation between GDM Risk with cumulation of the *HSPA1L* gene rs2227956 polymorphism, and gravidity.

These changes in odds were not statistically significant($P>0.05$).

7.1.4 Association of the *PPARG* gene variant rs3856806 and rs1801282 with Gravidity

Interaction analyses of the *PPARG* gene variants rs3856806 and rs1801282, gravidity, and GDM revealed remarkable changes in odds ratios under all genetic models adjusted for family history of diabetes (**Appendix A8: Table A23**). No increase in odds has resulted from any genotypes of the rs3856806 under any tested models within the primigravida group. Multigravidity itself increases the odds of having GDM in the presence of all genotypes under all tested models. Even the homozygous genotype of reference allele C increased the odds by 1.5 folds (OR=1.46) in the multigravida group under codominant and dominant models. The CT and TT genotypes resulted 1.6 (codominant and overdominant models) and 1.3 (codominant and recessive models) folds higher odds in multigravida group (**Figure 7.4**).

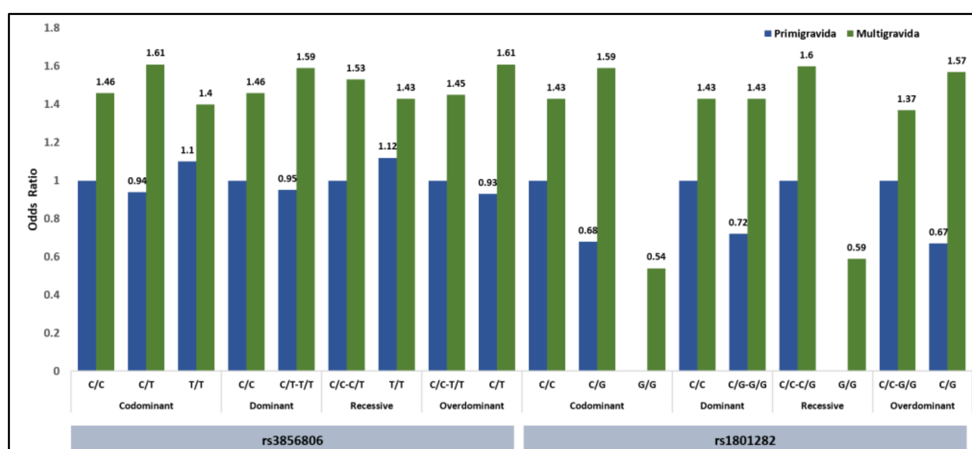


Figure 7.4: Association of GDM Risk with cumulation of the *PPARG* gene polymorphisms, and gravidity.

Within the primigravida control group, there was no individual with GG genotype. The CG genotype decreased the odds by 1.5 folds under codominant (OR=0.68) and overdominant (OR=0.67) models. Even the homozygous genotype of reference allele C increased the odds by 1.4 folds in the multigravida group under codominant and dominant models than those in primigravida group. The heterozygous genotype CG resulted 1.6 folds higher odds in multigravida group under codominant and overdominant models. None of these odds were statistically significant(P>0.05) (**Appendix A8: Table A24**).

7.1.5 Association of the *TCF7L2* gene variant rs10885406 with Gravidity

Interaction analyses of the *TCF7L2* gene variant rs10885406, gravidity, and GDM revealed no increase in odds under any tested models within the primigravida group; rather, GG genotype decrease odds by 1.7 folds under codominant (OR=0.60, 95% CI=0.22-1.63) and recessive (OR=0.59, 95% CI=0.23-1.48) models. Conversely, in multigravida group odds of having GDM increased compared to those in primigravida group. Even the reference allele A increased the odds by 1.5 folds under codominant (OR=1.48, 95% CI=0.81-2.68) and dominant (OR=1.47, 95% CI=0.81-2.67) models. The AG and GG genotype resulted 1.4- and 2.6-folds higher odds than primigravida group under tested models (**Appendix A8: Table A25**).

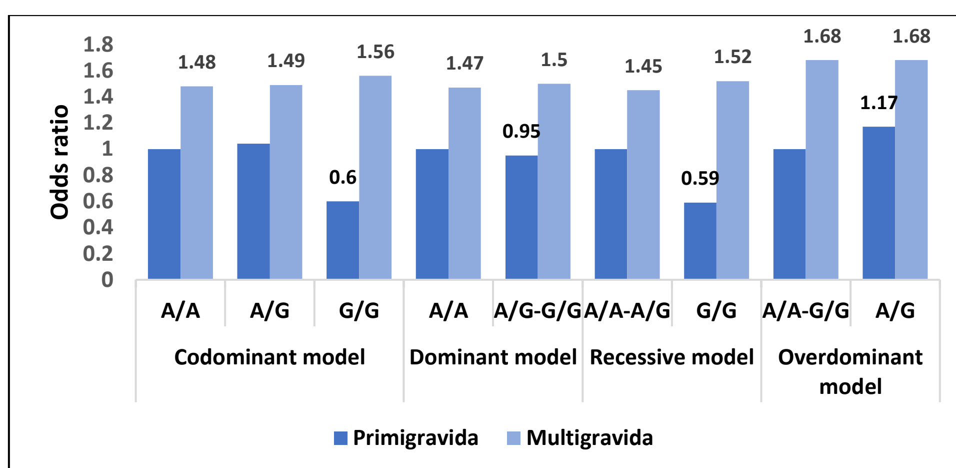


Figure 7.5: Correlation between GDM Risk with cumulation of the *TCF7L2* gene rs10885406 polymorphisms, and gravidity.

None of the changes in odds were statistically significant(P>0.05).

7.1.6 Association of the *TCF7L2* gene rs7903146 variant with Gravidity

Interaction analyses of the *TCF7L2* gene rs7903146 variant, gravidity, and GDM revealed that the TT genotype decrease in odds of having GDM by 2 folds under codominant (OR=0.48) and recessive (OR=0.50) models within the primigravida group. Conversely, the same genotype (TT) increased the odds in the multigravida group. The reference allele C increased the odds by 1.4 folds under codominant and dominant models than those in the primigravida group. The heterozygous genotype CT

increase the odds 1.5 folds from primigravida group under codominant (OR=1.34) and overdominant (OR=1.47) models (**Appendix A8: Table A26**).

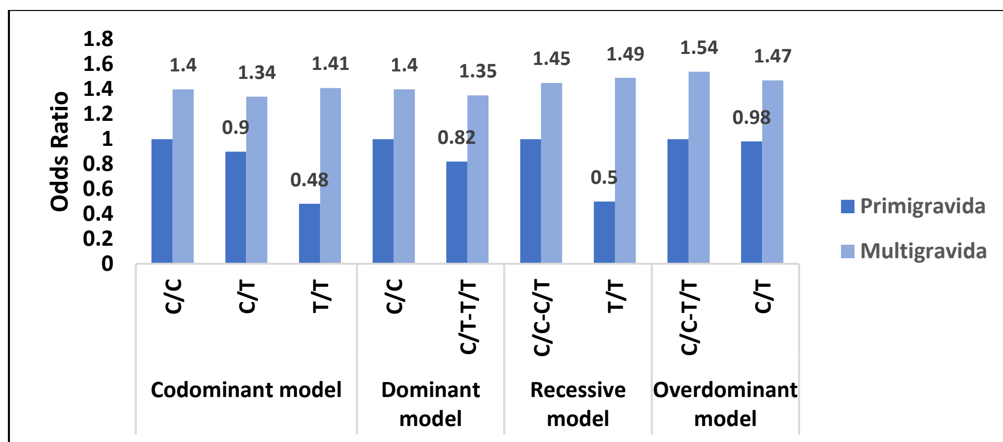


Figure 7.6: Association of GDM Risk with cumulation of the *TCF7L2* gene rs7903146 polymorphisms, and gravidity.

None of these changes in odds were statistically significant(P>0.05).

7.1.7 Association of the *TCF7L2* gene variant rs12255372 with Gravidity

Interaction analyses of the *TCF7L2* gene rs12255372 variant, gravidity, and GDM revealed no increase in odds from any genotypes under any tested models within the primigravida group. On the other hand, multigravid women had 1.2 to 2.5 folds increased in odds of having GDM. Even the reference allele G increased the odds by 1.2 folds under codominant and dominant models than those in the primigravida group.

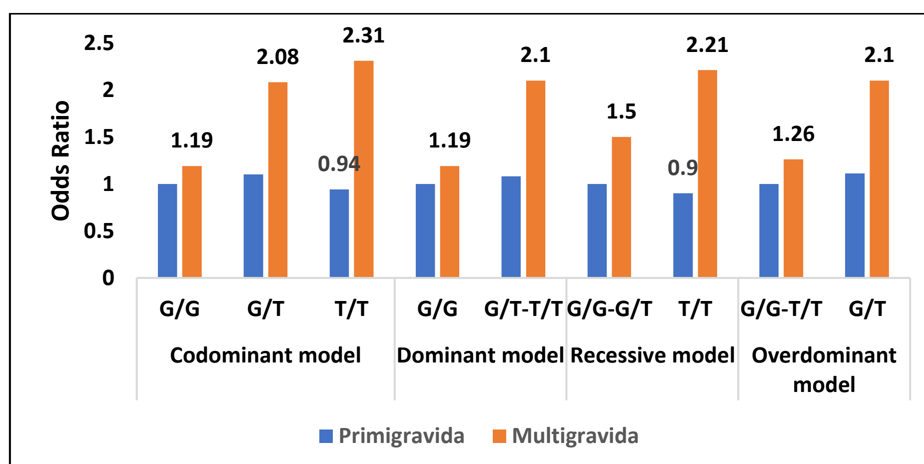


Figure 7.7: Interaction between GDM Risk with cumulation of the *TCF7L2* gene rs12255372 polymorphism, and gravidity.

None of these changes in odds were statistically significant(P>0.05) (**Appendix A8: Table A27**).

7.1.8 Association of the *WFS1* gene variant rs10010131 with Gravidity

Interaction analyses of the *WFS1* gene rs10010131 variant, gravidity, and GDM revealed a substantial increase in odds ratios under all genetic models adjusted for FHD. Within the primigravida group, the homozygous genotype of altered allele A increases the odds of GDM by 1.6 folds under codominant (OR=1.59) and recessive (OR=1.56) models. The reference allele G increased the odds by 1.8 folds in the multigravida group under codominant and dominant models than those in the primigravida group (**Appendix A8: Table A28**).

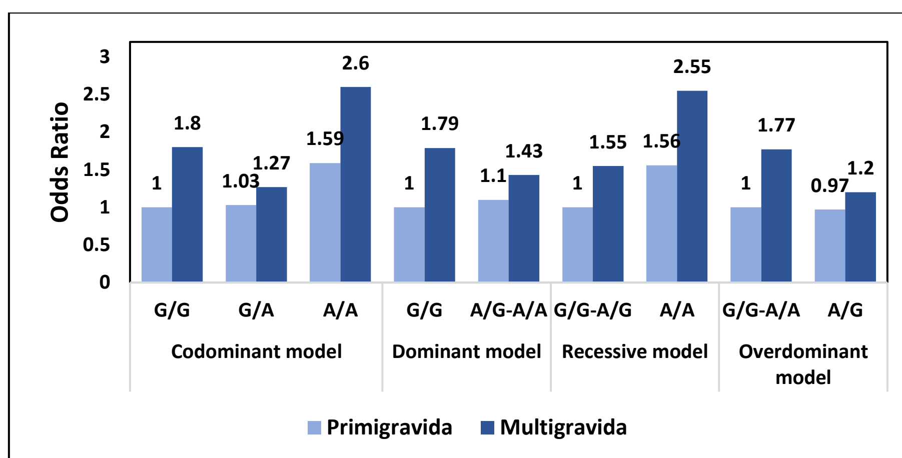


Figure 7.8: Cumulative association of GDM Risk with the *WFS1* gene variant rs10010131, and gravidity.

None of these changes in odds were statistically significant (P>0.05).

7.2 Summary

- The cumulative impact of target SNPs and gravidity on GDM were analyzed under **four genetic models**; codominant, dominant, recessive and overdominant models.
- The interaction analyses were adjusted for the family history of diabetes.
- Multigravidity and all genotypes (except CC genotype of the *FTO* gene rs8050136) of the target SNPs cumulatively increase odds of GDM under all tested models. Only the cumulative association of the *FTO* gene rs8050136 and

gravidity with GDM was statistically significant under codominant, dominant and overdominant models.

- The odds (except CC genotype of the *HSPA1L* gene rs2227956) of having GDM are higher in multigravida group than those in primigravida group. These differences were not statistically significant.

Reference

1. Zargar AH, Sheikh MI, Bashir MI, Masoodi SR, Laway BA, Wani AI, et al. Prevalence of gestational diabetes mellitus in Kashmiri women from the Indian subcontinent. Diabetes research and clinical practice. 2004;66(2):139-45.
2. Mierzyński R, Poniedziałek-Czajkowska E, Dłuski D, Kamiński M, Mierzyńska A, Leszczyńska-Gorzelak BJJoDR. The potential role of chemerin, lipocalin 2, and apelin in the diagnosis and pathophysiology of gestational diabetes mellitus. 2021;2021.
3. Rajput M, Bairwa M, Rajput R. Prevalence of gestational diabetes mellitus in rural Haryana: A community-based study. Indian J Endocrinol Metab. 2014;18(3):350-4.
4. Sivakumar V, Rajasekeran A, Vijayakumar AJIJPSR. Assessment of risk factors for the early detection of gestational diabetes mellitus. 2014;5(3):114-8.

8. Association of selected SNPs with anthropometric and metabolic parameters

8.1 Prevalence of GDM in the study participants

The prevalence of GDM in the study participants was calculated and was found to be 41.01%. The groups of participants with a positive family history of diabetes (52.28%) and multigravidity (47.42%) had a higher prevalence of this disease than the total participants (43.37%). Pregnant women recruited in the first trimester have a higher percentage (54.17%) of GDM cases than the total study participants and the participants recruited in the other two trimesters. The prevalence was also high in the higher age and BMI groups (**Table 8.1**).

Table 8.1: Prevalence estimates of GDM in different study characteristics

Variables		*Prevalence (%)*	*P value*
Family history of diabetes			
	No	37.62	
	Yes	52.28	0.0012
Bad obstetric History			
	No	43	
	Yes	44.36	0.7861
Occupation, number (%)			
	Housewife	43.79	
	Service Holder	49.12	0.32
	Other	2.8	
Gravidity			
	Primigravida	37.85	
	Multigravida	47.42	0.032
Trimester			
	First trimester	54.17	
	Second trimester	38.30	0.02
	Third trimester	42.92	0.09
Age strata (Years)			
	18-23	27.27	
	24-28	44.17	0.0013
	29-32	56.86	0.0361
	33-45	56.86	
BMI (kg/m^2)			
	≤ 18.5	28.57	
	18.6–24.9	32.26	0.7759
	25–29.9	48.33	0.0012
	≥30	54.55	0.3792

8.2 Fasting insulin level and insulin indices

This study was aimed to see the insulin indices in a part of the study samples comprising GDM (n=74) and NGT (n=87) and to compare these between the two groups. An independent sample Mann-Whitney U test revealed that fasting insulin levels were significantly higher in the GDM group (Median=9.52, n=74) compared to the control group (Median=6.71, n=87), U=2061, Z=3.93, r=0.3 with a P value less than 0.001.

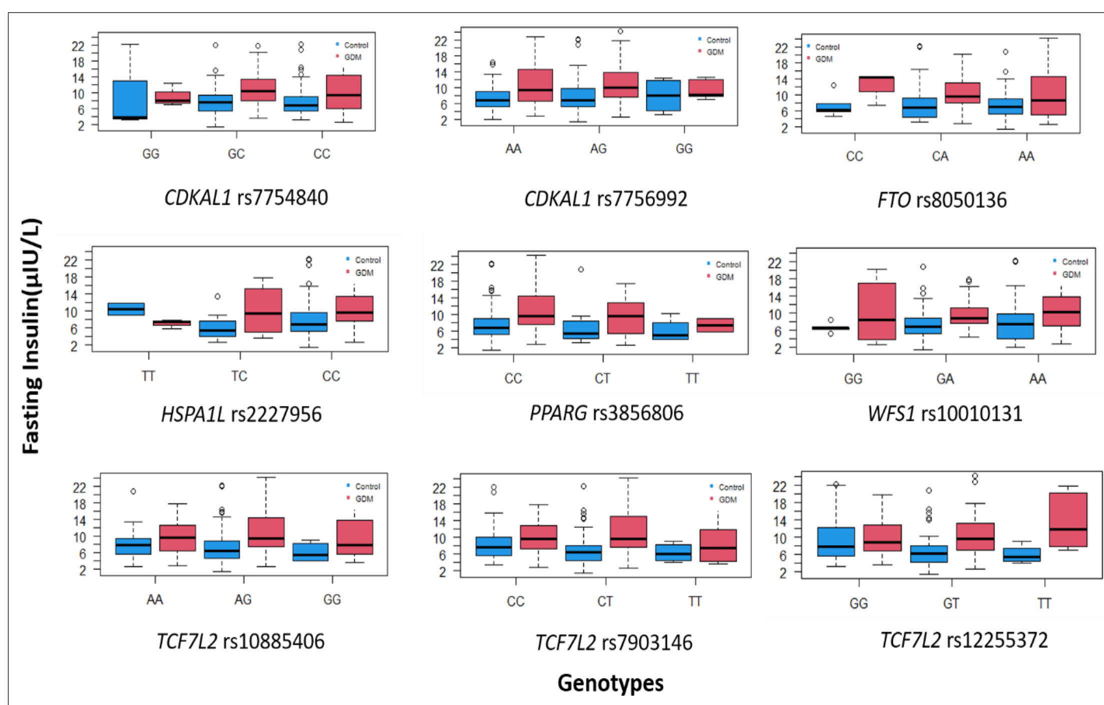


Figure 8.1: Fasting insulin levels between control and GDM groups in each genotype of the target SNPs

Fasting insulin levels between control and GDM groups in each genotype of the target SNPs are compared and shown the Figure 8.1. The *PPARG* gene variant rs1801282 was excluded from this comparison because of the absence of all three genotypes in control and GDM groups. The fasting insulin level was higher in GDM group than in control in all the genotypes for all of the SNPs except for the TT genotype in the *HSPA1L* rs2227956. In GDM group, the higher insulin level was observed in risk genotypes of the *CDKAL1, HSPA1L, TCF7L2* (rs12255372) and *WFS1* gene variants and reverse was observed in *TCF7L2* (rs10885406 and rs7903146), *FTO* and *PPARG* (rs3856806) variants.

Comparison of insulin indices between these two groups are shown in the **Table 8.2.** HOMA-IR was significantly higher in GDM than those of NGT whereas HOMA-B and

HOMA-%S were significantly lower in GDM than those of NGT. The differences of HOMA-IR and HOMA-%S between these two groups have large effect as r is greater than 0.5.

Table 8.2: Insulin indices in a part of study subjects

Variables	Median		*U*	*Z*	r	*P*
	Control (n = 86)	GDM (n = 73)				
HOMA-IR	1.30	2.33	1539	5.53	0.44	0.0001
HOMA-B	135.15	108.00	2141	3.44	0.3	0.001
HOMA-%S	77.24	60.96	1539	5.53	0.44	0.0001

Mann- Whitney U test for comparing groups; r=0.1, small effect; r=0.3, medium effect; r=0.5, large effect.

The association of the target SNPs with GDM with increasing fasting insulin, HOMA-IR, HOMA-B and HOMA-%S have been analyzed by using dominant (MM vs. Mm-mm) genetic model. Study subjects were divided into two groups for each of the parameters (fasting insulin, HOMA-IR, HOMA-B and HOMA-%S) and odds of GDM revealed from these two groups were compared. The comparisons are shown in Figure 8.2.

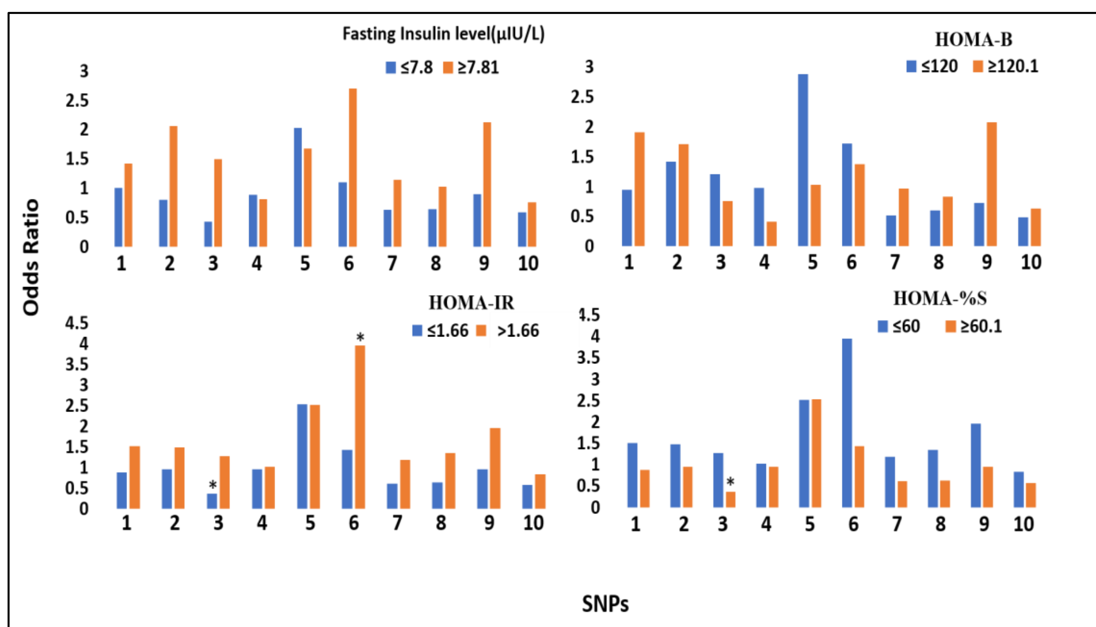


Figure 8.2: Fasting insulin levels between control and GDM groups in each genotype of the target SNPs;1: *CDKAL1* rs7754840; 2: *CDKAL1* rs7756992; 3: *FTO* rs8050136; 4: *HSPA1L* rs2227956; 5: *PPARG* rs3856806; 6: *PPARG* rs1801282; 7: *TCF7L2* rs10885406; 8: *TCF7L2* rs7903146; 9: *TCF7L2* rs12255372; 10: *WFS1* rs10010131.

The resulted odds of having GDM were higher in the high fasting insulin, HOMA-IR and HOMA-B groups except for the *HSPA1L* rs2227956, *TCF7L2* (rs10885406 and rs7903146), *PPARG* rs3856806 and *WFS1* rs10010131 variants. In high HOMA-IR group, *PPARG* rs1801282 shows significantly higher odds of having GDM and *FTO* rs8050136 shows statistically significant protection against GDM in low HOMA-IR and high HOMA-%S groups (Figure 8.2).

8.3 Trimester

Participants of this study were recruited irrespective of trimester. To determine the impact of trimester on anthropometric and demographic parameters, participants were divided into three groups according to their gestation weeks. The most of study subjects (46.14%) were in their third trimester, 37.23% in their second trimester, and the rest (14.26%) were in their first trimester. Among the total GDM diagnosed participants, 18% were in first trimester, 37% were in second trimester and 45% were in third trimester of pregnancy (**Figure 8.3**). Despite the fact that just about 15% of the study participants were in their first trimester, GDM diagnoses were nearly 1.3-1.4 times higher in the first trimester group than in the second and third trimester groups (**Table 8.1**).

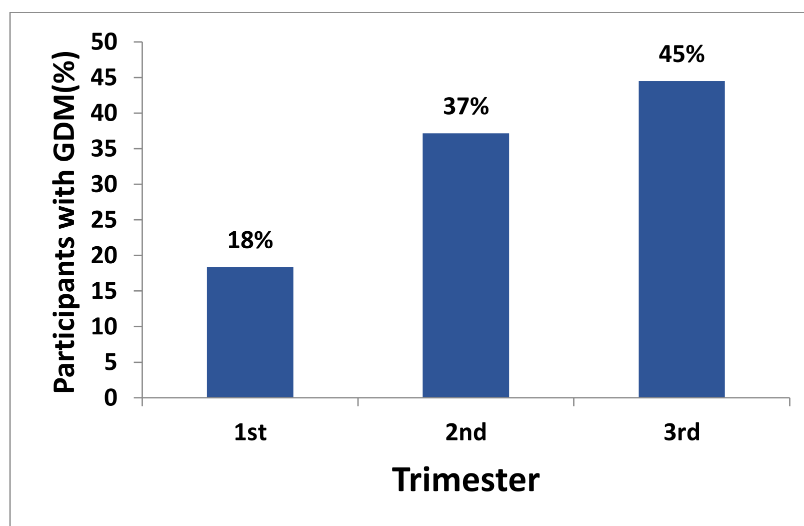


Figure 8.3: Percentage of total diagnosed GDM in different trimesters

One way ANOVA analyses revealed no significant differences of the variables between these three groups except the BMI and diastolic blood pressure of the normoglycemic

control individuals (**Appendix A9: Table A29**). Within each gestation group, the means of variables were compared between cases and controls (**Appendix A9: Table A30**). Plasma glucose levels were significantly higher in cases (GDM) than in controls within each trimester group. In 2nd and 3rd trimester, age and BMI of participants with GDM were significantly higher than that of controls (Figure 8.4)

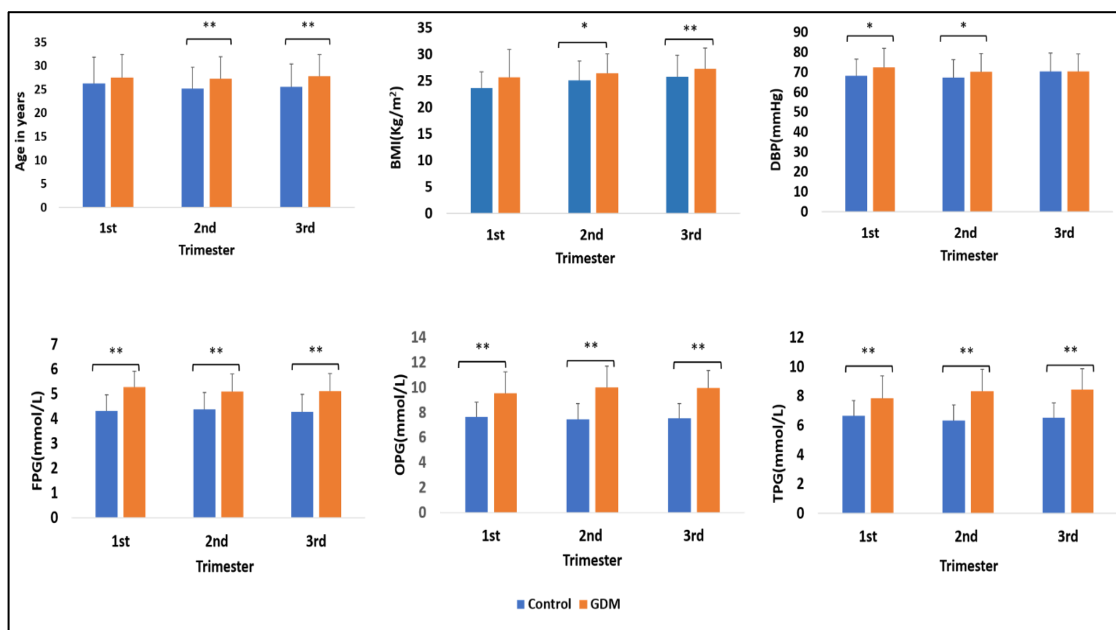


Figure 8.4: Comparison of study variables between cases and controls within three trimesters.

8.4 Age strata

The study participants were divided into four age groups (Table 8.1). A high prevalence (56.86%) of GDM was found in pregnant women aged between 29 – 32 and 33-45 age groups. In age groups 24-28 and 18-23 the prevalence of this disease was 44.17% and 27.27% respectively (Table 8.1). The associations of the target SNPs with GDM in each group have been analyzed and the results obtained for the *PPARG* rs1801282 and *TCF7L2* rs12255372 are shown in the figure 8.5. In case of *TCF7L2* variant rs12255372, odds of having GDM increased with the increasing age. In presence of the altered allele of the *PPARG* rs1801282 the risk of GDM decreased with the increasing age (Figure 8.5). In the remaining eight target SNPs, the rise in chances of GDM showed no such trends.

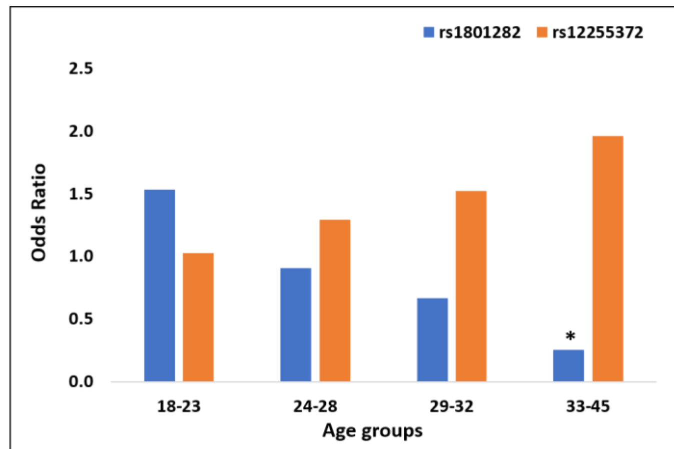


Figure 8.5: Association of SNPs with GDM in different age groups.

8.5 BMI Groups

The study participants were divided into four BMI group adopted according to the recommendations of the World Health Organization: <18.5, 18.5–24.9, 25.0–29.9, and ≥30 kg/m^2 for underweight, normal weight, overweight, and obesity, respectively (**Table 8.1**) and the associations of the target SNPs with GDM in each group have been analyzed. In case of the *CDKAL1* gene variants (rs7756992 and rs7754840) the odds of GDM increases with the increasing BMI whereas in case of *WFS1* rs10010131 these changes were in opposite direction (**Figure 8.6**). The increase in odds did not show such patterns in case of the rest of the target SNPs.

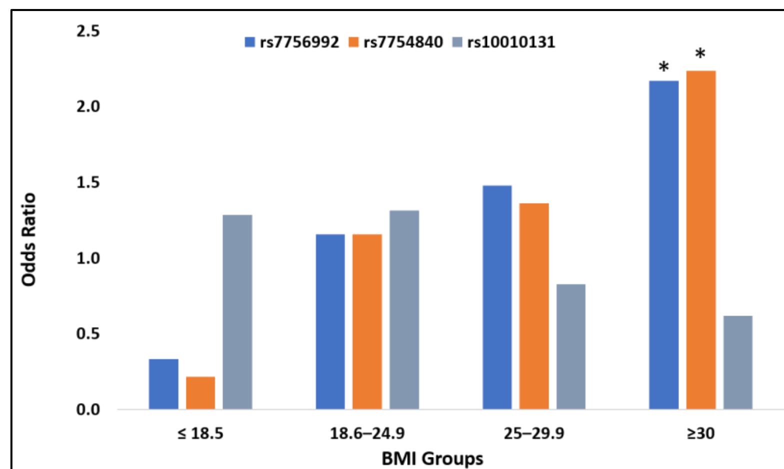


Figure 8.6: The changes in odds of GDM with the increase in BMI of the participants

The prevalence of GDM increased with increasing BMI and was highest in the group of participants with BMI ≥30 kg/m^2(**Table 8.1**). The association between minor alleles of the target SNPs and BMI (over 30 kg/m^2) have been tested to examine whether these alleles conferred a risk of GDM. The analyses revealed increased odds of having GDM by the minor alleles of the *CDKAL1*, *PPARG* and *HSPA1L* gene variants (**Figure 8.7**). The increases resulted from the association of the *CDKAL1* gene variants were statistically significant.

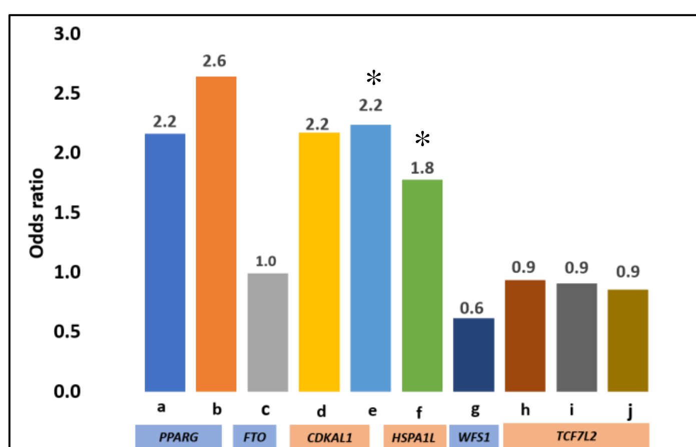


Figure 8.7: Association of the target SNPs with GDM within the BMI group ≥30 kg/m^2. a: rs3856806, b: rs1801282, c: rs8050136, d: rs7756992, e: rs7754840, f: rs2227956, g: rs10010131; h: rs12255372, i: rs7903146, j: rs10885406. (*, **, *** *P*<0.05, *P*<0.01, *P*<0.001)

8.6 Summary

- The prevalence of GDM in the studied population was 41.01%.
- Prevalence of GDM was high in 1st trimester group.
- High prevalence of GDM was observed in group of participants with multigravida, positive family history of diabetes, older age and high BMI.
- Fasting insulin level and HOMA-IR were significantly higher in GDM group.
- HOMA-B and HOMA-%S were significantly higher in control group.
- The odds of GDM changes with the increase of age and BMI. These changes were associated with the genotypes of the target SNPs

9. Discussion

GDM is an ever-increasing complication of pregnancy. Women with GDM is prone to have greater than 7-fold risk of T2DM development in later life (1). In this study, we intended to examine the association between T2DM-related SNPs and the predisposition to GDM in a sample of Bangladeshi women. On the basis of the literature review, ten SNPs of six genes were selected, among which eight are intron variants and two are coding sequence variants (**Section 3.1**). Although the majority of disease-associated SNPs are identified in exons or coding regions, there is evidence that SNPs can also occur in intronic regions of genes, altering the regulatory region and thereby affecting the splicing mechanism and gene expression (2). A total of 534 pregnant women irrespective of trimester were screened for GDM, among whom 505, comprising 286 normoglycemic control and 219 GDM, were finally selected as study subjects (**Section 4.1**). Genotyping of the target SNPs in the study subjects were carried out by PCR-RFLP, T-ARMS and TaqMan allelic discrimination assay methods (**Section 3.9**). The association of the selected SNPs with GDM was analyzed by using five genetic models adjusted for confounding variables (**Section 3.12**).

9.1 Prevalence of GDM

GDM prevalence has increased rapidly in Bangladesh (3-5). We screened 534 Bangladeshi expecting women, and 219 of them were diagnosed with gestational diabetes, i.e., the prevalence of this disease in our study subject was 41.01% (Section 8.1). The incidence of this disease was increased by 1.2% compared with previously reported data (35% to 41.01%) (5). Like previous reports (6-8) women with gravida ≥2 and positive family history of diabetes of this study had a significantly higher prevalence of GDM (**Table 8.1**). In our study, prevalence in different trimester was much higher **(Table 8.1)** (9). Such high prevalence of GDM in Bangladesh could be contributed by several factors. In addition, for diagnosis of GDM the most recent WHO criteria has been used, with a lower FPG threshold value (5.1–6.9 mmol/L) (10) in comparison with the earlier studies done in Bangladesh that used either FPG ≥7.0 mmol/L or FPG ≥5.3 mmol/L as cutoff points (3, 11-13). This differential prevalence of GDM is reported in different studies globally since the diagnostic criteria for GDM are different and an increased rate was obtained with the updated diagnostic criterion (14).

9.2 Screening and Diagnosis of GDM

Confirming the GDM cases was crucial as well as challenging in the current study. As the pregnancy advances, the placental hormone mediated insulin resistance increases. Therefore, early testing may not be beneficial for diagnosing GDM. Insulin resistance increases throughout the second trimester (13-28 weeks of pregnancy), and hence the level of glucose rises in those women who cannot produce enough insulin to cope with this resistance. Moreover, carrying out the test too late in the 3rd trimester ultimately confines the time of metabolic interventions. So, screening for GDM is preferable at 24-28 weeks of gestation (15). Samples were collected regardless of trimester, and women who had normal plasma glucose levels earlier in pregnancy were tested again at the 24-28th week of pregnancy for confirmation. Moreover, the common practice in our country is to perform a two-sample OGTT, i.e., fasting and 2 hr after glucose load; the plasma glucose levels are measured for GDM detection. In our study, we followed the WHO 2013 criteria where in addition to these two-time points, plasma glucose levels were measured after 1 hr of the glucose load. Twenty individuals (9.13% of the total diagnosed) were diagnosed with GDM based on 1hr plasma glucose level who otherwise remained untreated and unmanaged.

9.3 Genotype and allele distribution of the target SNPs

In the current study, the genotype distribution of the target SNP followed HWE in the control group. This means, absence of selection bias, genotyping error as well as absence of population stratification in the study population (16). The cases need to be evaluated as well in addition to the controls, to avoid eliminating important potential causal SNPs of a common disease (17). Hence, we included the assessment of the distribution of genotypes of the target SNPs in cases as well for HWE. Significant departure from the equilibrium was found for the *CDKAL1* gene variant rs7756992 (**Table 4.3**) and the *TCF7L2* gene variant rs12255372 (**Table 4.9**) which indicates a possible association of this SNP with GDM.

The G allele frequency (0.12) of *PPARG* rs1801282 in the GDM group was lowest among the studied SNPs and was similar to the Danish (0.12), and greater than those reported for the Arab (0.07), French (0.10), Greek (0.02), Korean (0.05), and Chinese

(0.05) populations, whereas Scandinavian (0.15) and Turkish (0.19) populations showed a higher frequency of the G allele (18-21). Regional differences in population genetics can explain these findings.

In our study, MAF for rs7756992 (G allele) and rs12255372 (T allele) were 0.328 and 0.28. In contrast, the MAFs reported in 1000 Genomes were 0.269 and 0.22 in South Asian populations. The disparity between our MAF and the MAF published by 1000 Genomes could be due to genetic variations between our group and the study populations used by 1000 Genomes. The minor allele frequencies of the remaining eight SNPs were comparable with that reported in 1000 Genomes for South Asian populations.

9.4 Selection of Confounder variables

Any variable must have three criteria to be considered as a potential confounder (22). First is to be differently distributed between the case and control groups, i.e., must be associated with the disease. Secondly, it must not be an effect of the disease or be a factor in the causal pathway of the disease and finally, it should be a risk factor for the disease. Systolic blood pressure, bad obstetric history, and occupational (Housewife and Service Holder) history were comparable between the control and cases. In contrast, significant differences were observed in age, BMI, Diastolic Blood Pressure, occupational (Others) history, gravidity, and family history of diabetes (**Table 4.1**). As a result, the latter variables were potential confounders. The association of these variables with GDM risk was analyzed by multivariate logistic regression, and odds ratios are shown in **Table 5.1**. Risk factors were determined from the value of odds ratios. The higher risk resulted from the associations of gravidity (**OR=1.5**) and family history of diabetes (**OR=1.84**) with GDM. Consequently, these are the confounders that have been adjusted for in the subsequent association analyses. Confounding, unlike other biases, can be controlled by employing stratification to account for it once a study is completed (22). As a result, the participants in this study were divided into strata or subgroups based on the levels of confounding factors. Cross-classification interactions were used to conduct relative risk analysis for each stratum in the control and GDM groups using different genetic models. (**Appendix A7, A8**).

9.5 Association of T2DM related SNPs with GDM

Among others, the six selected genes (*CDKAL1*, *FTO*, *HSPA1L*, *PPARG*, *TCF7L2* and *WFS1*) of this study are the members of first set of T2DM susceptibility genes identified by GWA and other studies (23-31). Variants of these genes reported to be related with T2DM in Asian, European, and American populations had previously been discovered (23, 24, 32-41). Because this disease may share a genetic background with T2DM, the connection of these genetic variants with GDM was investigated in a number of populations (42-44). Recently some studies on these variants reported significant association in Asian populations (33, 45-47) whereas in some populations there was no association found (44, 48). Despite this, there is no report published on the association of these variants with GDM in Bangladeshi women except for *TCF7L2* rs7903146 (49).

Our data also demonstrates a significant association of *CDKAL1* rs7756992, *CDKAL1* rs7754840, and *TCF7L2* rs12255372 with GDM, which is consistent with the positive associations reported in the Asian population. Our findings on association of *CDKAL1* variants with GDM have been published (51).

9.5.1 Association of *CDKAL1* gene variants

The effect size of the rs7756992 G allele (OR =1.36), in this study is close to the other reported studies for Asians (OR = 1.41), and South Indians (OR = 1.45) (**Table 5.2**). In the current study, several genetic models were used for the association analyses e.g., codominant, dominant, recessive, overdominant, and log additive. This could avoid possible biases in identification of significant associations (52). Significant association between this SNP and GDM was revealed by the codominant (P=0.047), dominant (P=0.014) and log additive (P=0.018) models (**Appendix A6: Table A1**). The association of rs7756992 with GDM was also significant under dominant and log additive models despite adjusting for family history of diabetes and gravidity. The GG genotype showed a higher odd ratio (1.81) after adjustment and the odd of AG genotype remained comparable (**Figure 5.1**).

In our study, there was no significant association observed between rs7754840 and GDM risk before adjustment for confounders. Despite, the homozygous risk allele of this SNP showed higher odds of having GDM in comparison with that of rs7756992

(**Appendix A6: Table A2**). However, the recessive model showed a significant association (*P*=0.047) between the variant and GDM after adjustment for family history of diabetes and gravidity (**Figure 5.2**). Our findings are consistent with studies in Caucasian, Korean and South Indian populations, which showed a good and significant associations (32, 50, 54) on the other hand no significant association was observed in a Chinese and Egyptian study (44, 53). Differences in ethnicities, diagnostic criteria for GDM and finally the sample size can be the reason behind these variations in association.

9.5.2 Association of *FTO* gene variant

In congruence with Cho et al. (OR=1.12, *P*=0.30) and Saucedo, Renata et al. (OR=1.11, *P*=0.86) we found no significant association (*P*=0.18) between the *FTO* rs8050136 with GDM but odds (OR=1.63) of this disease is 1.5 folds higher in our studied population (55, 56). Not only the odds of GDM but also the minor allele frequency of this variant is much higher (0.291) than that of Korean (0.122) and Mexican (0.138) populations. These findings indicate the risk providing nature of this SNP.

9.5.3 Association of *HSPA1L* gene variant

The *HSPA1L* gene rs2227956 has been reported to be associated with diabetic Nephropathy, sarcoidosis and T2DM (57-59). To our best knowledge this is the first study to find out association of this SNP with GDM. In our study this polymorphism did not show any effect on susceptibility to GDM under any genetic model (**Figure 5.4 and Appendix A6: Table A4**). In our study, the frequency of the TT genotype of *HSPA1L* (*Hsp70-hom*) as well as the frequencies of the C and T allele were equal to the frequency observed in normoglycemic pregnant mothers and these may describe the cause of no association (**Table 4.5**).

9.5.4 Association of *PPARG* gene variants

A number of groups reported association of *PPARG* rs3856806 with T2DM susceptibility. But no association (OR=1.03) of this variant with GDM was observed in our study. This finding is similar (OR=1.09) to that of the Korean population (55) (**Figure 5.5 and Appendix A6: Table A5**). No difference was observed between

frequencies of CC, CT and TT genotypes and alleles of *PPARG* rs3856806 in cases and controls of our study (**Table 4.6**) and these outcomes in genotype and allele frequencies may explain why the association was not found.

The *PPARG* rs1801282 polymorphism results in a substitution of proline for alanine at codon 12 of exon B (C > G; Pro12Ala). This polymorphism causes a conformational change in the protein, and the presence of the minor allele is associated with a reduction in the activity of PPARG (60). The association of this SNP with T2DM as well as GDM is contested. Positive associations have been reported by some studies while others showed protection against or no association with these disease entities (19, 42, 61). No association of this variant with GDM has been reported in Korean (32), Scandinavian, or Arab pregnant women (20). The GG genotype of *PPARG* rs1801282 showed protection (OR<1) against GDM in this studied population under codominant (OR=0.57) and recessive (OR=0.58) models with a *P* value >0.05 (**Figure 5.6).**

9.5.5 Association of *TCF7L2* gene variants

The *TCF7L2* gene has been regarded as the most common susceptible gene for T2DM among various ethnic groups in the world (62). However, different races and ethnicities have varied genetic backgrounds, including risk allele frequency and SNP linkage disequilibrium distribution (63, 64). It is the first study of *TCF7L2* variants rs12255372 and rs10885406 in our population. Regarding rs12255372 SNP, GT and TT genotypes were more prominent among the GDM group compare to the control group and significantly increased the odds of having GDM by 1.4-fold under dominant model (**Figure 5.9**). Women carrying the T allele have a nearly 1.3-fold increase risk for GDM (**Table 5.2**). Similar results have been reported in other studies (65, 66).

For *TCF7L2* variant rs10885406, the odds of GDM also showed protective nature (OR<1) under codominant and recessive models (**Figure 5.7**). This observation was not statistically significant (**Appendix A6: Table A7**).

The *TCF7L2* rs7903146 changed allele is linked to a loss of initial postprandial glycemic control, emphasizing the significance of insulin treatment in pregnant women with chronic hyperglycemia (67). In contrast to a previous study with *TCF7L2* rs7903146 (49), our analyses revealed protective (OR = 0.72) nature of the altered

genotypes against GDM (**Figure 5.8 and Appendix A6: Table A8**). This discrepancy can be explained by the difference in sample size of the two studies (505 vs. 100). Findings of our study are in accordance with a recent study, where the T allele of this variant lower the risk of GDM (68).

9.5.6 Association of *WFS1* gene variant

In our studied population, the risk allele A of this variant was not associated (OR=1) with GDM (**Table 5.2**) but the AA genotype increase odds of GDM by 1.61 folds under recessive model. This result is in contrast to the outcome (OR 0.87) reported in Lauenborg et al. (42)

Our finding is that the risk alleles for the *PPARG* rs1801282, *TCF7L2* rs10885406, and rs7903146 can be identified as "protective" alleles (**Table 5.7**). Contrasting results such as those in this study are common in genetic studies (69) and maybe attributed to population unique linkage disequilibrium patterns, or that the identified polymorphism is likely not the causal variant (70).

9.6 Association analysis of haplotypes with GDM

The target SNPs that were located in the same chromosome have been analyzed for the association of their haplotypes with GDM. The haplotype analysis of *CDKAL1* variants rs7756992 and rs7754840 (**Table 5.3**) identified a significant association of haplotype of both risk alleles (CG) and GDM susceptibility proving the association. The *CDKAL1* and *HSPA1L* gene also located on the same chromosome (**Table 3.1**) and we analyzed the association of their haplotypes with GDM. The risk alleles containing haplotype (GCC) of these three SNPs increase odds of GDM by 2.77 folds (**Table 5.4**). Whereas the haplotypes of *CDKAL1* variants increase the odds by 1.43-folds (**Table 5.3**). Moreover, the risk allele of *HSPA1L* variant did not show any effect on GDM (**Section 5.1.3**) but in combination with *CDKAL1* variants increased the risk of GDM.

The association of haplotypes of *PPARG* gene variants revealed that the risk alleles containing haplotype (TG) of this gene provide protection against GDM (**Table 5.5**); otherwise, the rs3856806 has no effect on this disease (**Section 5.1.4**). The same

association for the haplotypes of *TCF7L2* gene variants did not show any effect on GDM (**Table 5.6**).

9.7 Cumulative association of the target SNPs with GDM and family history of diabetes

We have investigated confirmed type 2 diabetes-associated SNPs for association with GDM. Yet, we have failed to find evidence of association between GDM and these confirmed loci, with the exception of the *CDKAL1* gene variants, rs7756992 and rs7754840, and the *TCF7L2* gene variant rs12255372. Family history of diabetes is well reported risk factor for GDM, T2DM and T2D related complications (71-73). In our study it was also detected as a confounding variable (**Section 5.1**). Moreover, one of the specific objectives of this study is to observe the association of tested polymorphisms in GDM women with positive family history of T2DM. For that reason, we analyzed the cumulative impact of target SNPs and family history of diabetes on GDM. In case of all SNPs, there were substantial increases in the odds of GDM except for *HSPAL1* rs2227956 which provide protection against GDM in presence of family history of diabetes (**Section 6.1.3**).

Under the dominant genetic model, the combined effect of the CDKAL1 rs7756992 polymorphism and a positive family history of diabetes elevated the risk of GDM by 2.7-fold, which was significant (**Figure 6.1**). The AG and GG genotypes of this SNP increased the risk of GDM by 2.5 and 3.7 times, respectively, although none of these genotypes were found to be harmful in those who had no family history of diabetes (**Figure 6.1**). An association between the CC genotype of the rs7754840 polymorphism with late-onset T2DM patients was suggested by Mansoori et al (74). In the current study, the CC genotype of the SNP increased the GDM odds 4.5 times and the GG genotype increased it 1.72 times in women with previous history of diabetes in the family (**Figure 6.2**). The reference allele of *TCF7L2* gene variant rs10885406 significantly increase the odds of GDM by 3.4 folds in presence of family history of diabetes under codominant and dominant models (**Figure 6.6**) whereas this variant provide protection against GDM (**Table 5.7**).

9.8 Synergetic association of the target SNPs with GDM and gravidity

As gravidity is one of the confounders of our study the interaction of the target SNPs, gravidity and GDM have been carried out and adjusted for family history of diabetes. In primigravida we did not find any risk of GDM provided by the altered alleles of the studied gene variants. The *TCF7L2* rs7903146 and rs10885406, *PPARG* rs1801282 and *FTO* rs8050136 showed protection against GDM.

Interaction analyses revealed substantial increases in odds of GDM in the multigravida group by all the target SNPs. Even the reference alleles of these variants showed notable increase in the risk of this disease. But these increases were not statistically significant except for *FTO* rs8050136. Although *FTO* did not show any significant association with GDM (**Section 5.1.2**), in presence of multigravidity the risk allele of *FTO* rs8050136 significantly increase the odds of having GDM under codominant, dominant and overdominant models (**Figure 7.2**)

The wider CI values and insignificant *P* values of different outcomes can be caused by small sample size of the subgroups of both confounders (**Appendix A7 and A8**).

9.9 General characteristics of the study population

Other studies in our country observed significantly higher mean age and BMI in GDM in comparison to NGT mothers (3, 75). The present study also noted older age and significantly higher BMI in GDM mothers compared to NGT. Family history of diabetes was found to be higher in GDM which is also in agreement with previous studies (49, 75).

9.9.1 Age

While some studies (76, 77), have suggested that GDM risk increases linearly with maternal age, another (78) showed that the incidence of GDM increased with age, highest at 35–39 years and then declined in women aged 40–50 years. When the associations of the target SNPs with GDM have been analyzed in our study, it revealed that for *TCF7L2* rs12255372, odds of having GDM increased with the increasing age

(**Figure 8.5**). For *PPARG* rs1801282 the odds decreased with increasing age and at 33-45 years age groups this SNP provide significant protection against GDM (**Figure 8.5**). Furthermore, Zhang *et al*. (79) showed the highest age-specific prevalence of GDM in 30–34 years old women. In our study we found increase in prevalence of GDM with increasing age (**Table 8.1**) and peaked at 29-45 years.

9.9.2 BMI

BMI is widely considered as a confounder of T2DM as well as GDM risk. An increase in BMI can cause diabetes with severe risk factors like raised insulin, glucose, and triglycerides. BMI was not the risk factor for the study samples as odds of GDM was 1.04 with a *P* value <0.001 **(Table 5.1)**. Based on stratified analysis, high prevalence (32.26%) of GDM was resulted even in the normal weight group. Association analysis of the target SNPs with GDM in different BMI groups revealed that for *WFS1* rs10010131 odds of GDM decreased with increasing BMI but for *CDKAL1*gene variants the odds increased with increasing BMI (**Figure 8.6**). In the BMI group ≥30 kg/m^2, *FTO* rs8050136 has no association (OR=1) with GDM (**Figure 8.6**) whereas this SNP increase the odds of GDM by 1.63 folds (**Table 5.7**).

9.9.3 Gestation weeks

The study participants were recruited irrespective of trimester. The study participants were divided into three groups according to their gestation weeks. Ifat Ara Begum reported 12.5% of the total diagnosed GDM in the study subjects were in first trimester, 31.2% were in second trimester and 56.3 % were in third trimester of pregnancy (80). In our study subjects 18% of the total diagnosed GDM are in first trimester, 37% were in second trimester and 45% were in third trimester of pregnancy (**Figure 8.3**). The differences in the percentages may explained by the difference in the sample size (117 vs. 505) and type of participants (Primigravida vs. both primi- and multigravida) of these two studies. In a previous study trimester wise prevalence of GDM was reported as 1st vs. 2nd vs. 3rd trimester: 40.9% vs. 44.9% vs. 48.7% (3). In our study population trimester wise prevalence are 54.17%, 38.30% and 42.92% respectively (**Table 8.1**).

The prevalence in 1st trimester is much higher in our study. This supports the method of trimester independent sample collection (**Section 3.2.1**).

9.9.4 Fasting Insulin level and insulin indices

Insulin resistance in peripheral tissues is a high-risk factor and almost invariably precedes the development of overt type 2 diabetes (81). Although pregnancy is a condition characterized by progressive insulin resistance, GDM develops in only a small proportion of pregnant women (82). Insulin resistance was assessed by indirect measures using HOMA model as the gold standard clamp method could not be availed (**Section 3.4**). Statistically significant between-group differences were observed for fasting insulin level and insulin indices. Insulin resistance measured as HOMA-IR was significantly higher and insulin secretory capacity measured as HOMA-B, insulin sensitivity measured as HOMA-%S were significantly lower in GDM than that of NGT (**Table 8.2**).

It is notable that the SNPs that associated with increased GDM risk were also observed to increase the risk of higher fasting insulin levels (**Figure 8.2**). However, interpretation is complex, as (83) fasting insulin is strongly influenced by insulin sensitivity (84). The changes in odds by *PPARG* rs3856806 and *WFS1* rs10010131 variants could not be explained by this. The rs1801282 C allele is associated with increased transcriptional activity of PPARG and, consequently, increased sensitivity to insulin. This may explain the significant increase in odds of having GDM by *PPARG* rs1801282 G allele in high HOMA-IR group (**Figure 8.2**). The significant protection against GDM provided by *FTO* rs8050136 in low HOMA-IR and high HOMA-%S groups can be explained by the hypothesis that *FTO* rs8050136 may affect HOMA-IR directly or indirectly. It's plausible to assume that people who possess the risk allele of FTO are predisposed to insulin resistance and can boost insulin production to entirely compensate for low insulin sensitivity in order to maintain glucose homeostasis.

9.10 Power of the study

GAS power calculator (85) was used to detect the power of study and considering the total sample size (both cases and controls) studied, genotype relative risks for the target

SNPs, the prevalence of GDM (0.41) in the study participants, disease allele frequency obtained, the significance level equal to 0.05 we had power greater than 0.8 for the target SNPs. The study power equal to 0.8 or greater indicates that the study sample size had sufficient power to detect the association.

9.11 The strength and limitations of this study

In our study all participants are of the same ethnic origin which is a strong point of the study. Moreover, the subjects were analyzed in a standardized manner where the criteria for diagnosis was well defined. The genotyping was performed blind concerning case control status.

The current study's findings might be specific for the inhabitant of Dhaka city GDM patients where the city lifestyle might have an effect as the individuals were recruited from Dhaka city and nearby regions like Narayanganj and Gazipur. This also indicates the nature of the association of the target SNPs and GDM in the current study. Despite considering several risk factors in the study, there are some other GDM associated factors which were left out, e.g food intake and physical activity which could not be considered due to the unavailability of information. In addition, fasting insulin levels and the insulin indices that could be risk factors for GDM in our study were only measured in a part of the total study samples. Hence, the interactions between these genetic variants and lifestyle remain unrevealed. Furthermore, eight selected SNPs were identified to be located in the intron region; and the relationships between these SNPs and genes and their mechanism behind modulation of GDM risk are largely unknown.

Reference

1. Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet (London, England). 2009;373(9677):1773-9.

2. Datta A, Mazumder MHH, Chowdhury AS, Hasan MA. Functional and Structural Consequences of Damaging Single Nucleotide Polymorphisms in Human Prostate Cancer Predisposition Gene RNASEL. Biomed Res Int. 2015;2015:271458-.

3. Nusrat-Sultana HM, Sharmin-Jahan M-H, Sandesh-Panthi Y-A, Fariduddin MJStJ. Alarming frequency of Gestational Diabetes Mellitus (GDM) attending a tertiary care hospital in Bangladesh. 2016.
4. Sandesh-Panthi MA, Hasanat M-H, Yasmin-Aktar N-S, Sharmin-Jahan MFJJ. Frequency of gestational diabetes mellitus in Bangladesh impact of WHO 2013 screening criteria: Efficiency DIPSI & WHO 1999 criteria. 2015;2(2):13-9.
5. Mazumder T, Akter E, Rahman SM, Islam MT, Talukder MR. Prevalence and Risk Factors of Gestational Diabetes Mellitus in Bangladesh: Findings from Demographic Health Survey 2017-2018. Int J Environ Res Public Health. 2022;19(5):2583.
6. Rajput M, Bairwa M, Rajput R. Prevalence of gestational diabetes mellitus in rural Haryana: A community-based study. Indian J Endocrinol Metab. 2014;18(3):350-4.
7. Garshasbi A, Faghihzadeh S, Naghizadeh MM, Ghavam MJJoF, Health R. Prevalence and risk factors for gestational diabetes mellitus in Tehran. 2008:75-80.
8. Mwanri AW, Kinabo J, Ramaiya K, Feskens EJJDr, practice c. Prevalence of gestational diabetes mellitus in urban and rural Tanzania. 2014;103(1):71-8.
9. Seshiah V, Balaji V, Balaji MS, Paneerselvam A, Arthi T, Thamizharasi M, et al. Gestational diabetes mellitus manifests in all trimesters of pregnancy. Diabetes research and clinical practice. 2007;77(3):482-4.
10. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: a World Health Organization Guideline. Diabetes research and clinical practice. 2014;103(3):341-63.
11. Mannan MA, Rahman MH, Ara I, Afroz H. Prevalence and Pregnancy Outcome of Gestational Diabetes Mellitus Among Bangladeshi Urban Pregnant Women. Journal of Medicine. 2012;13(2):147-51.
12. Sayeed MA, Jahan S, Rhaman MM, Chowdhury MMH, Khanam PA, Begum T, et al. Prevalence and perinatal outcomes in GDM and non-GDM in a rural pregnancy cohort of Bangladesh. Ibrahim Medical College Journal. 2014;7(2):21-7.
13. Jesmin S, Akter S, Akashi H, Al-Mamun A, Rahman MA, Islam MM, et al. Screening for gestational diabetes mellitus and its prevalence in Bangladesh. Diabetes research and clinical practice. 2014;103(1):57-62.

14. Behboudi-Gandevani S, Amiri M, Bidhendi Yarandi R, Ramezani Tehrani F. The impact of diagnostic criteria for gestational diabetes on its prevalence: a systematic review and meta-analysis. Diabetology & Metabolic Syndrome. 2019;11(1):11.

15. Diabetes IAo, care PSGCPJD. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. 2010;33(3):676-82.

16. Namipashaki A, Razaghi-Moghadam Z, Ansari-Pour NJCJ. The essentiality of reporting Hardy-Weinberg equilibrium calculations in population-based genetic association studies. 2015;17(2):187.

17. Wang J, Shete SJHh. Using both cases and controls for testing hardy-weinberg proportions in a genetic association study. 2010;69(3):212-8.

18. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316(5826):889-94.

19. Pappa KI, Gazouli M, Economou K, Daskalakis G, Anastasiou E, Anagnou NP, et al. Gestational diabetes mellitus shares polymorphisms of genes associated with insulin resistance and type 2 diabetes in the Greek population. Gynecological endocrinology : the official journal of the International Society of Gynecological Endocrinology. 2011;27(4):267-72.

20. Shaat N, Ekelund M, Lernmark A, Ivarsson S, Nilsson A, Perfekt R, et al. Genotypic and phenotypic differences between Arabian and Scandinavian women with gestational diabetes mellitus. Diabetologia. 2004;47(5):878-84.

21. Santos IC, Frigeri HR, Réa RR, Almeida AC, Souza EM, Pedrosa FO, et al. The glucokinase gene promoter polymorphism -30G>A (rs1799884) is associated with fasting glucose in healthy pregnant women but not with gestational diabetes. Clinica chimica acta; international journal of clinical chemistry. 2010;411(11-12):892-3.

22. Jager K, Zoccali C, Macleod A, Dekker FJKi. Confounding: what it is and how to deal with it. 2008;73(3):256-60.

23. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. 2007;39(6):770-5.

24. Dehwah M, Wang M, Huang QJGMR. CDKAL1 and type 2 diabetes: a global meta-analysis. 2010;9(2):1109-20.

25. Duesing K, Fatemifar G, Charpentier G, Marre M, Tichet J, Hercberg S, et al. Strong association of common variants in the CDKN2A/CDKN2B region with type 2 diabetes in French Europids. 2008;51(5):821-6.

26. Hertel J, Johansson S, Raeder H, Midthjell K, Lyssenko V, Groop L, et al. Genetic analysis of recently identified type 2 diabetes loci in 1,638 unselected patients with type 2 diabetes and 1,858 control participants from a Norwegian population-based cohort (the HUNT study). 2008;51(6):971-7.

27. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. 2007;316(5829):1331-6.

28. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. 2007;316(5829):1341-5.

29. Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, Ohnaka K, et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. 2009;58(7):1690-9.

30. Basile KJ, Johnson ME, Xia Q, Grant SFJIjoe. Genetic susceptibility to type 2 diabetes and obesity: follow-up of findings from genome-wide association studies. 2014;2014.

31. Doria A, Patti M-E, Kahn CRJCm. The emerging genetic architecture of type 2 diabetes. 2008;8(3):186-200.

32. Cho Y, Kim T, Lim S, Choi S, Shin H, Lee H, et al. Type 2 diabetes-associated genetic variants discovered in the recent genome-wide association studies are related to gestational diabetes mellitus in the Korean population. 2009;52(2):253-61.

33. Kwak SH, Kim SH, Cho YM, Go MJ, Cho YS, Choi SH, et al. A genome-wide association study of gestational diabetes mellitus in Korean women. Diabetes. 2012;61(2):531-41.

34. Rong R, Hanson RL, Ortiz D, Wiedrich C, Kobes S, Knowler WC, et al. Association analysis of variation in/near FTO, CDKAL1, SLC30A8, HHEX, EXT2,

IGF2BP2, LOC387761, and CDKN2B with type 2 diabetes and related quantitative traits in Pima Indians. 2009;58(2):478-88.

35. Hu C, Zhang R, Wang C, Wang J, Ma X, Lu J, et al. PPARG, KCNJ11, CDKAL1, CDKN2A-CDKN2B, IDE-KIF11-HHEX, IGF2BP2 and SLC30A8 are associated with type 2 diabetes in a chinese population. 2009;4(10):e7643.

36. Ng MC, Park KS, Oh B, Tam CH, Cho YM, Shin HD, et al. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. Diabetes. 2008;57(8):2226-33.

37. Omori S, Tanaka Y, Takahashi A, Hirose H, Kashiwagi A, Kaku K, et al. Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8, and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. 2008;57(3):791-5.

38. Wu Y, Li H, Loos RJ, Yu Z, Ye X, Chen L, et al. Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. 2008;57(10):2834-42.

39. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. 2008;40(5):638-45.

40. Moniruzzaman M, Ahmed I, Huq S, All Mahmud MS, Begum S, Mahzabin Amin U, et al. Association of polymorphism in heat shock protein 70 genes with type 2 diabetes in Bangladeshi population. 2020;8(2):e1073.

41. Pagán A, Sabater-Molina M, Olza J, Prieto-Sánchez MT, Blanco-Carnero JE, Parrilla JJ, et al. A gene variant in the transcription factor 7-like 2 (TCF7L2) is associated with an increased risk of gestational diabetes mellitus. 2014;180:77-82.

42. Lauenborg J, Grarup N, Damm P, Borch-Johnsen K, Jørgensen T, Pedersen O, et al. Common type 2 diabetes risk gene variants associate with gestational diabetes. 2009;94(1):145-50.

43. Tarnowski M, Malinowski D, Pawlak K, Dziedziejko V, Safranow K, Pawlik AJCJoD. GCK, GCKR, FADS1, DGKB/TMEM195 and CDKAL1 gene polymorphisms in women with gestational diabetes. 2017;41(4):372-9.

44. Wang Y, Nie M, Li W, Ping F, Hu Y, Ma L, et al. Association of six single nucleotide polymorphisms with gestational diabetes mellitus in a Chinese population. 2011;6(11):e26953.
45. Cho YM, Kim TH, Lim S, Choi SH, Shin HD, Lee HK, et al. Type 2 diabetes-associated genetic variants discovered in the recent genome-wide association studies are related to gestational diabetes mellitus in the Korean population. Diabetologia. 2009;52(2):253-61.
46. Guo F, Long W, Zhou W, Zhang B, Liu J, Yu BJAoG, et al. FTO, GCKR, CDKAL1 and CDKN2A/B gene polymorphisms and the risk of gestational diabetes mellitus: a meta-analysis. 2018;298(4):705-15.
47. Zhang C, Bao W, Rong Y, Yang H, Bowers K, Yeung E, et al. Genetic variants and the risk of gestational diabetes mellitus: a systematic review. 2013;19(4):376-90.
48. Xie K, Chen T, Zhang Y, Wen J, Cui X, You L, et al. Association of rs10830962 polymorphism with gestational diabetes mellitus risk in a Chinese population. 2019;9(1):1-6.
49. Hasan M, Hasnat MA, Hasan KN, Panthi S, A Y, Sultana N, et al. TCF7L2 gene rs7903146 polymorphism is observed in gestational diabetes mellitus in Bangladesh. Integrative Obesity and Diabetes. 2016.
50. Kanthimathi S, Chidambaram M, Liju S, Bhavadharini B, Bodhini D, Prakash VG, et al. Identification of genetic variants of gestational diabetes in South Indians. 2015;17(7):462-7.
51. Amin USM, Parvez N, Rahman TA, Hasan MR, Das KC, Jahan S, et al. CDKAL1 gene rs7756992 A/G and rs7754840 G/C polymorphisms are associated with gestational diabetes mellitus in a sample of Bangladeshi population: implication for future T2DM prophylaxis. Diabetol Metab Syndr. 2022;14(1):18.
52. Salanti G, Southam L, Altshuler D, Ardlie K, Barroso I, Boehnke M, et al. Underlying genetic models of inheritance in established type 2 diabetes associations. 2009;170(5):537-45.
53. El Noury A, Azmy O, Alsharnoubi J, Salama S, Okasha A, Gouda WJBrn. Variants of CDKAL1 rs7754840 (G/C) and CDKN2A/2B rs10811661 (C/T) with gestational diabetes: insignificant association. 2018;11(1):1-6.

54. Rosta K, Al-Aissa Z, Hadarits O, Harreiter J, Nádasdi Á, Kelemen F, et al. Association study with 77 SNPs confirms the robust role for the rs10830963/G of MTNR1B variant and identifies two novel associations in gestational diabetes mellitus development. 2017;12(1):e0169781.

55. Cho YM, Kim TH, Lim S, Choi SH, Shin HD, Lee HK, et al. Type 2 diabetes-associated genetic variants discovered in the recent genome-wide association studies are related to gestational diabetes mellitus in the Korean population. Diabetologia. 2009;52(2):253-61.

56. Saucedo R, Valencia J, Gutierrez C, Basurto L, Hernandez M, Puello E, et al. Gene variants in the FTO gene are associated with adiponectin and TNF-alpha levels in gestational diabetes mellitus. Diabetology & metabolic syndrome. 2017;9:32-.

57. Dhamodharan U, Ezhilarasi K, Ponjayanthi B, Sireesh D, Ramkumar KM, Viswanathan V. Association of A1538G and C2437T single nucleotide polymorphisms in heat shock protein-70 genes with diabetic nephropathy among South Indian population. Biosci Rep. 2017;37(2):BSR20160605.

58. Bogunia-Kubik K, Koscinska K, Suchnicki K, Lange A. HSP70-hom gene single nucleotide (+2763 G/A and +2437 C/T) polymorphisms in sarcoidosis. International journal of immunogenetics. 2006;33:135-40.

59. Umapathy D, Krishnamoorthy E, Muthukumaran PJJoDFC. Is HSP70-hom (C2437T) Single Nucleotide Polymorphism (SNP) asso-ciated with Diabetic Foot Ulcer (DFU) among South Indian population? 2012;4:57-62.

60. Meirhaeghe A, Amouyel P. Impact of genetic variation of PPARgamma in humans. Molecular genetics and metabolism. 2004;83(1-2):93-102.

61. Groop L, Pociot F. Genetics of diabetes--are we missing the genes or the disease? Molecular and cellular endocrinology. 2014;382(1):726-39.

62. Fritsche L, Sarief M, Wagner R, Stefan N, Lehmann R, Häring HU, et al. Genetic variation in TCF7L2 rs7903146 and history of GDM negatively and independently impact on diabetes-associated metabolic traits. Diabetes research and clinical practice. 2018;146:251-7.

63. Luo Y, Wang H, Han X, Ren Q, Wang F, Zhang X, et al. Meta-analysis of the association between SNPs in TCF7L2 and type 2 diabetes in East Asian population. 2009;85(2):139-46.

64. Qiao H, Zhang X, Zhao X, Zhao Y, Xu L, Sun H, et al. Genetic variants of TCF7L2 are associated with type 2 diabetes in a northeastern Chinese population. 2012;495(2):115-9.
65. Watanabe RM, Allayee H, Xiang AH, Trigo E, Hartiala J, Lawrence JM, et al. Transcription factor 7-like 2 (TCF7L2) is associated with gestational diabetes mellitus and interacts with adiposity to alter insulin secretion in Mexican Americans. Diabetes. 2007;56(5):1481-5.
66. Francaite-Daugeliene M, Lesauskaite V, Tamosiunas A, Jasukaitiene A, Velickienė D. Genetic variants of TCF7L2 gene and its coherence with metabolic parameters in Lithuanian (Kaunas district) women population with previously diagnosed gestational diabetes mellitus compared to general population. Diabetes research and clinical practice. 2021;172:108636.
67. Potasso L, Perakakis N, Lamprinou A, Polyzou E, Kassanos D, Peter A, et al. Clinical Impact of the TCF7L2 Gene rs7903146 Type 2 Diabetes Mellitus Risk Polymorphism in Women with Gestational Diabetes Mellitus: Impaired Glycemic Control and Increased Need of Insulin Therapy. Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association. 2020;128(10):663-6.
68. Barabash A, Valerio JD, Garcia de la Torre N, Jimenez I, Del Valle L, Melero V, et al. TCF7L2 rs7903146 polymorphism modulates the association between adherence to a Mediterranean diet and the risk of gestational diabetes mellitus. Metabolism open. 2020;8:100069.
69. Lin PI, Vance JM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. American journal of human genetics. 2007;80(3):531-8.
70. Klimentidis YC, Lemas DJ, Wiener HH, O'Brien DM, Havel PJ, Stanhope KL, et al. CDKAL1 and HHEX are associated with type 2 diabetes-related traits among Yup'ik people. Journal of diabetes. 2014;6(3):251-9.
71. Maghbooli Z, Pasalar P, Keshtkar A, Farzadfar F, Larijani BJJod, disorders m. Predictive factors of diabetic complications: a possible link between family history of diabetes and diabetic retinopathy. 2014;13(1):1-5.

72. Xiong X-f, Wei L, Xiao Y, Han Y-C, Yang J, Zhao H, et al. Family history of diabetes is associated with diabetic foot complications in type 2 diabetes. Sci Rep. 2020;10(1):17056.

73. Moosazadeh M, Asemi Z, Lankarani KB, Tabrizi R, Maharlouei N, Naghibzadeh-Tahami A, et al. Family history of diabetes and the risk of gestational diabetes mellitus in Iran: A systematic review and meta-analysis. Diabetes & metabolic syndrome. 2017;11 Suppl 1:S99-s104.

74. Mansoori Y, Daraei A, Naghizadeh MM, Salehi RJAbr. Significance of a common variant in the CDKAL1 gene with susceptibility to type 2 diabetes mellitus in Iranian population. 2015;4.

75. Panthi S, Hasanat M, Hasan M, Aktar Y, Sultana N, Jahan S, et al. Frequency of Gestational Diabetes Mellitus in Bangladesh: Impact of WHO 2013 Screening Criteria : Efficiency of DIPSI and WHO 1999 Criteria JOURNAL OF CLINICAL DIABETOLOGY. Journal of Clinical Diabetology. 2015;2:13-20.

76. Liu X, Zhang WJCmj. Effect of maternal age on pregnancy: a retrospective cohort study. 2014;127(12):2241-6.

77. Schummers L, Hutcheon JA, Hacker MR, VanderWeele TJ, Williams PL, McElrath TA, et al. Absolute risks of obstetric outcomes risks by maternal age at first birth: a population-based cohort. 2018;29(3):379.

78. Wang Y, Chen L, Xiao K, Horswell R, Besse J, Johnson J, et al. Increasing incidence of gestational diabetes mellitus in Louisiana, 1997–2009. 2012;21(3):319-25.

79. Zhang F, Dong L, Zhang C, Li B, Wen J, Gao W, et al. Increasing prevalence of gestational diabetes mellitus in Chinese women from 1999 to 2008. 2011;28(6):652-7.

80. Begum I. Gestational Diabetes Mellitus in Primi Gravida of Bangladesh in Different Trimesters. International Journal of Biology. 2014;6.

81. Østergård T, Ek J, Hamid Y, Saltin B, Pedersen OB, Hansen T, et al. Influence of the PPAR-gamma2 Pro12Ala and ACE I/D polymorphisms on insulin sensitivity and training effects in healthy offspring of type 2 diabetic subjects. Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme. 2005;37(2):99-105.

82. Buchanan TA, Xiang AH. Gestational diabetes mellitus. The Journal of clinical investigation. 2005;115(3):485-91.

83. Almind K, Delahaye L, Hansen T, Van Obberghen E, Pedersen O, Kahn CRJPotNAoS. Characterization of the Met326Ile variant of phosphatidylinositol 3-kinase p85α. 2002;99(4):2124-8.

84. Alpha B, Cox L, Crowther N, Clark PM, Hales C. Sensitive amplified immunoenzymometric assays (IEMA) for human insulin and intact proinsulin. 1992.

85. Johnson JL, Abecasis GR. GAS Power Calculator: web-based power calculator for genetic association studies. 2017:164343.

10.Conclusion

10.1 Summary

This is the first extensive association study on GDM in Bangladesh. This study revealed the importance of the diagnosis of GDM irrespective of trimester and following three steps diagnostic criteria. The prevalence of GDM in the studied population was 41.01%. High prevalence of GDM was observed in group of participants with multigravida, positive family history of diabetes, older age and high BMI. Fasting insulin level and HOMA-IR were significantly higher in GDM group. In high HOMA-IR group, *PPARG* rs1801282 showed significantly higher odds of having GDM and *FTO* rs8050136 shows statistically significant protection against GDM in low HOMA-IR and high HOMA-%S groups.

This study shows that CDKAL1 rs7756992 and rs7754840, as well as TCF7L2 rs12255372, are significantly associated with and raised the risk of GDM in a sample of pregnant Bangladeshi women, stressing the importance of these possibly functional polymorphisms in the development of GDM. When compared to rs7756992 and rs12255372, rs7754840 is not only strongly related with but also leads in a higher risk of GDM. For the identification of causative loci and genes, the underlying mechanisms must be clarified by functional investigations. *HSPAL1* rs2227956 and *PPARG* rs3856806 have no association with GDM in our sample of population. Although the sample size in this analysis provided enough power to detect significant relationships between GDM and chosen variants, stratification for confounding factors reduced the number of samples in each stratum, resulting in negligible associations with variations in some strata. In presence of family history of diabetes all genotypes of the target SNPs substantially increased (1.5 to 4 folds) the odds of GDM only except the homozygous risk allele carrier genotype of *HSPA1L* variant. Significant susceptible association with GDM in presence of positive family history of diabetes was revealed only from the interaction analyses of the *CDKAL1* rs7756992 and the *TCF7L2* rs10885406. Multigravidity and all genotypes (except CC genotype of the *FTO* gene rs8050136) of the target SNPs cumulatively increase odds of GDM under all tested models. Only the cumulative association of the *FTO* gene rs8050136 and gravidity with GDM was statistically significant.

These genetic predispositions of the investigated variants could be a valuable marker for identifying pregnant women at a higher risk of GDM, who could subsequently be exposed to earlier regular GDM screening as well as lifestyle management before and after pregnancy to avoid GDM and T2DM, respectively. Furthermore, modern research avenues in the search for cost-effective technology for routine early GDM screening will open up. In addition, substantial research with various ethnic communities is required to confirm our findings.

10.2 Importance in terms of public health

Prevention and early detection of GDM are critical for public health. The link between T2D-related SNPs and GDM predisposition may raise public awareness, particularly among women and health care professionals. To reduce the risk of GDM, effective interventions are needed to maintain a healthy lifestyle and metabolic condition in reproductive-age women. Furthermore, screening for GDM risk throughout the periconceptional period and the first month of pregnancy is required to allow for early interventions to manage hyperglycemia, hyperinsulinemia, and other associated metabolic abnormalities in order to avoid unfavorable fetal outcomes. GDM and T2D-related SNPs in mothers may suggest that GDM is a risk factor for later maternal T2DM. These women may be monitored and treated on a regular basis to help prevent diabetes in the long run. Furthermore, the relationship between these SNPs and the risk of GDM before or during conception deserves more exploration. Because the biological function of detected variants is still difficult to interpret, and people may carry the effects of one or more variants in several genes, the association between genetic variants and GDM alone, as well as gene-environment interaction effects and risk of T2DM, may be difficult to apply in public health. Approaches to dealing with many causative genes, as well as their successful application, are important in GDM genetic research. These links do not indicate to a single main T2D-related gene linked to GDM, but they do support the idea that T2D is causally linked to GDM through several T2D susceptibility gene variations and interactions with environmental variables. T2D-related SNPs, family history, and multigravidity were all linked to the risk of GDM in this dissertation. Mothers with GDM were more likely to develop metabolic problems later in life. This dissertation identified potential risk factors that can be used to prevent

GDM. This dissertation also included biomarkers that could be used to predict GDM. Furthermore, mothers with GDM must be monitored regularly for metabolic anomalies later in life.

10.3 Future Plan

This dissertation investigates the link between T2DM-related SNPs and GDM propensity. More research is needed to corroborate the findings and to overcome the study's shortcomings. This dissertation has a problem in that the data from participants is incomplete. We discovered that some maternal medical problems were underreported, as well as some missing data on possible confounders. Furthermore, the limited sample size leads to broad confidence intervals and effects of SNPs on the link between confounders (family history of diabetes and gravidity) and GDM risk. As a result, higher sample sizes are needed in future investigations to confirm these correlations. This research discovered links between T2DM-related SNPs and GDM. The FTO gene variation and multigravidity were revealed to have strong cumulative correlations with GDM. As a result, a prospective study that continuously measures plasma glucose levels from before conception to the succeeding pregnancy could help assess the prediction of GDM in multigravida women. Furthermore, to validate and compare the relationships between CDKAL1 gene variations and the risk of GDM, more research is needed with bigger sample sizes and other ethnic groups. Because many SNPs in the same gene were linked to GDM, haplotype analysis and gene-gene interactions may uncover additional areas or genes with small effects individually but evidence of linkage and association when combined. One of the study's weaknesses is that it didn't look into gene-environment interaction connections between T2DM genes and maternal environmental variables. Observational studies show a link between exposures and diseases, but they often fail to pinpoint the etiological mechanisms of disease caused by exposure. As a result, the substantial connections between T2D-related SNPs and diabetes family history on the risk of GDM necessitate additional investigation into the etiological mechanisms. Other studies will be needed to learn how diabetes in the family and multigravidity affect GDM occurrence. This study can be used as a pilot study for larger investigations in the future, such as constructing genome-wide association studies with next-generation sequencing, developing a strategy for

identifying SNPs for therapeutic use, or validating the related SNPs as biomarkers for T2DM and GDM. The results of this study will be confirmed after 5–6 years of follow-up.

Appendix

A1.

Data collection Sheet

ID (code) no: ------------------ Date: ------------------ Cross ID no: -------------------

Name of the Hospital: ……………………………………………………………………..

Name of Respondent: ………………………………………. Age: ………. Sex: Female

Occupation: ……………………………………………………………………………………

Home address: ………………………………………………..

……………………………………………………………………...

Telephone numbers: Mobile: …………………………...... Land: …………………………

Presenting Features:

Blood Pressure (BP) in mmHg: Systolic BP: [ ] Diastolic BP: [ ]

BMI (kg/m^2): [ ]

Menstrual history: Last Menstrual Period: [ ]

Obstetrical History:

Para : [ ] Gravida: [ ]

Gestation: [ ] weeks MR/Abortion: Yes / No

History of DM/GDM : Yes / No

Family history of T2DM among 1st degree relatives:

Parents: ………………..Siblings:…………………..Offsprings:

Any clinically significant illness: Yes/No

Result of Plasma glucose by OGTT: Status: GDM/Normal (any criteria)

Type of OGTT	Date & wk of gestation	Plasma glucose value (mmol/L)			Status
Time of glucose sample		0 Hr	01 Hr	02 Hr	
75 gram OGTT					
75 gram OGTT at 24 - 28th week of gestation, if –ve before 24 weeks					

Signature of researcher with Date

A2. DNA Extraction by Purelink™ Genomic DNA Mini Kit

Materials Needed for the extraction

- 96-100% ethanol
- Sample for DNA isolation
- Sterile, DNase-free microcentrifuge tubes
- Water baths or heat blocks and a microcentrifuge capable of centrifuging >10,000 x g

Preparation of Blood Lysate

1. A water bath or heat block was set at 55°C.
2. 20 µl Proteinase K was added to a sterile microcentrifuge tube.
3. Blood samples processing:
    - To a sterile microcentrifuge tube, added up to 200 µl fresh or frozen blood sample.

4. PBS was transferred to the tube containing Proteinase K from Step 2.
5. 20 µl RNase A was added to the sample. Mixed well by brief vortexing and incubated at room temperature for 2 minutes.
6. 200 µl PureLink™ Genomic Lysis/Binding Buffer was added and mixed well by vortexing to obtain a homogenous solution.
7. Incubated at 55°C for 10 minutes to promote protein digestion.
8. 200 µl 96-100% ethanol was added to the lysate. Mixed well by vortexing to yield a homogenous solution.

DNA Purification Protocol using Purelink™ Kit

1. A PureLink™ Spin Column was removed in a Collection Tube from the package.
2. The lysate (~640 µl) prepared with PureLink™ Genomic Lysis/Binding Buffer and ethanol was added to the spin column.
3. The column was centrifuged at 10,000 × g for 1 minute at room temperature.
4. The collection tube was discarded and placed the spin column into a clean PureLink™ Collection Tube supplied with the kit.
5. 500 µl Wash Buffer 1 prepared with ethanol was added to the column.
6. Column was centrifuged at 10,000 × g for 1 minute at room temperature.
7. The collection tube was discarded and placed the spin column into a clean PureLink™

collection tube supplied with the kit.

8. 500 µl Wash Buffer 2 prepared with ethanol was added to the column.

9. The column was centrifuged at maximum speed for 3 minutes at room temperature. Collection tube was discarded.

10. The spin column was placed in a sterile 1.5-ml microcentrifuge tube.

11. 50 µl of PureLink™ Genomic Elution Buffer was added to the column.

12. Incubated at room temperature for 1 minute. The column was centrifuged at maximum speed for 1 minute at room temperature.

13. To recover more DNA, a second elution step was performed using the same elution buffer volume as first elution.

14. The column was centrifuged at maximum speed for 1.5 minutes at room temperature. The tube contained purified DNA. Removed and discarded the column.

15. The purified DNA was stored at 4ºC (short-term) or -20°C (long-term) for downstream application.

A3. Agarose gel electrophoresis

A3.1 Materials

I. Ultrapure agarose (typing grade)

II. Electrophoresis buffer (TAE)

III. Gel loading buffer (6X)

IV. Ethidium bromide

V. DNA size standards/ DNA marker

VI. Deionized water

VII. Equipment for agarose gel electrophoresis:

a. Clean dry horizontal electrophoresis apparatus with chamber

b. Clean dry glass/ plastic plates with appropriate comb.

c. Gel-sealing tape

d. Power supply device

e. UV illuminator.

A3.2 Preparation of 0.8% agarose gel solution (100 ml)

- 0.8 g of agarose powder was weighed in a conical flask.
- 2 ml of 50× TAE buffer was taken in a measuring cylinder and the volume was made up to 100 ml with ddH2O and was poured in the flask containing Agarose and then melted in microoven at 60°C for 2 minutes.

Procedure

1. The open ends of a clean, dry plastic tray was sealed with tape and placed on a horizontal section of the bench.

2. Sufficient electrophoresis buffer (usually 1×TAE or 0.5×TBE) was prepared to fill the electrophoresis tank and to cast the gel.

3. A solution of agarose in electrophoresis buffer was prepared for separating the particular size fragments expected in the DNA sample.

4. The flask was placed in the microwave oven on high temperature (i.e.,60°C-80°C) for 2-3 minutes or until the agarose dissolves.

5. Using insulated gloves; the flask was transferred into a water bath at55°C. The gel solution was mixed thoroughly by gentle swirling when the molted gel had cooled.

6. The cooled gel solution was poured into the gel tray. An appropriate comb was placed \ previously for forming the sample slots in the gel and assuring that there were no bubbles around the combs. (A pipette tip was used to remove if there was any bubble.)

7. The gel was allowed to set completely (30-45 minutes at room temperature), then poured a small amount electrophoresis buffer on the top of the gel, and the comb was removed carefully. The electrophoresis buffer was poured off and the tape was removed carefully. The gel was mounted in the electrophoresis tank.

8. Electrophoresis buffer (1×TAE) was adjusted sufficiently to cover the gel to a depth of ~ 1 mm.

9. The samples of DNA was mixed with gel loading buffer (5:1) and loaded slowly into the slots of the submerged gel using a disposable micropipette. Size standard that will depend on the type of marker being analyzed was loaded into slots on both the right and left sides of the gel.

10. The lid of the gel tank was closed and the electrical leads were attached to the power supply device so that the DNA will migrate toward the positive anode (red lead).

11. A voltage of 1-5 V/cm (measured as the distance between the positive and negative electrodes) was applied.

12. The electric current was turn off when the DNA samples or dyes had migrated a sufficient distance through the gel and the leads and lid from the gel tank was removed.

13. The gel was stained by immersing it in electrophoresis buffer or H_2O containing ethidium bromide (0.5µg/ml) for 30-45 minutes at room temperature.

14. Photograph of the gel was taken under UV illumination.

A3.3 Preparation and maintenance of reagent for Agarose gel electrophoresis

i) Electrophoresis buffer (TAE buffer)

Components:

Stock solution (50X) Per Liter

 Tris base 242 g

 Glacial acetic acid 57.1

 0.5 M EDTA (pH 8.0) 100 ml

Working Solution (1×)

 Tris acetate 40 mM

 EDTA 1 mM

Procedure:

1. To prepare stock solution, all components were dissolved and then final volume was adjusted to 1000 ml with ddH_2O. Finally, it was sterilized by autoclaving.

2. 1 L/ 1000 ml of IX TAE working electrophoresis buffer was prepared by mixing 20 mlof 50 ×TAE stock solution and 980 ml of dd H_2O.

ii) Gel loading buffer (6X)

a. Sequencing dye

Components and amounts of sequencing dye

 Bromophenol blue 0.25%

 Xylene cyanol FF 0.25%

 Glycerol, in H2O 30%

This buffer was stored of at 4°C.

b. DNA dye

Components and amounts of DNA dye

 Bromophenol blue 0.05%

 Sucrose 40.0%

 EDTA 0.1M

 SDS 0.5%

iii) Ethidium bromide: Ethidium bromide was prepared as a stock solution of 10 mg/ml in H_2O, was stored at room temperature in dark bottles.

iv) DNA size standards/ DNA marker: A stock solution of size standards was prepared by dilution with a gel-loading buffer and TE (1:5:4) and then used as needed in individual Electrophoresis experiments.

A4. Composition, Preparation and Maintenance of PAGE

1. Acrylamide/ bis-acrylamide (40%)

Components of 40% acrylamide solution

Component	Amount
Acrylamide	190 g
N, N′-Methylene bis acrylamide	10 g

ddH2O, volume to	500 ml

The solution was heated to dissolve the chemicals. After preparing 40% acrylamide/bis acrylamide solution, it was filtered through Whatman filter paper and was stored in a dark bottle at 4°C.

2. Ammonium per sulfate (APS) (10%)

Components of 10% APS

Component	Amount
Ammonium per sulfate	1 g
ddH_2O, volume to	10 ml

After aliquoting in eppendorf tubes, the ammonium per sulfate (250 µl in each) was stored at –20°C.

3. TBE Buffer (stock solution) (5X)

Components of TBE buffer

Component	Amount
Tris base	54g
Boric acid	27.5 g
0.5 M EDTA (pH 8.0)	20 ml
ddH2O, volume to	1000 ml

N.B. The pH of the concentrated stock buffer should be ~8.3.

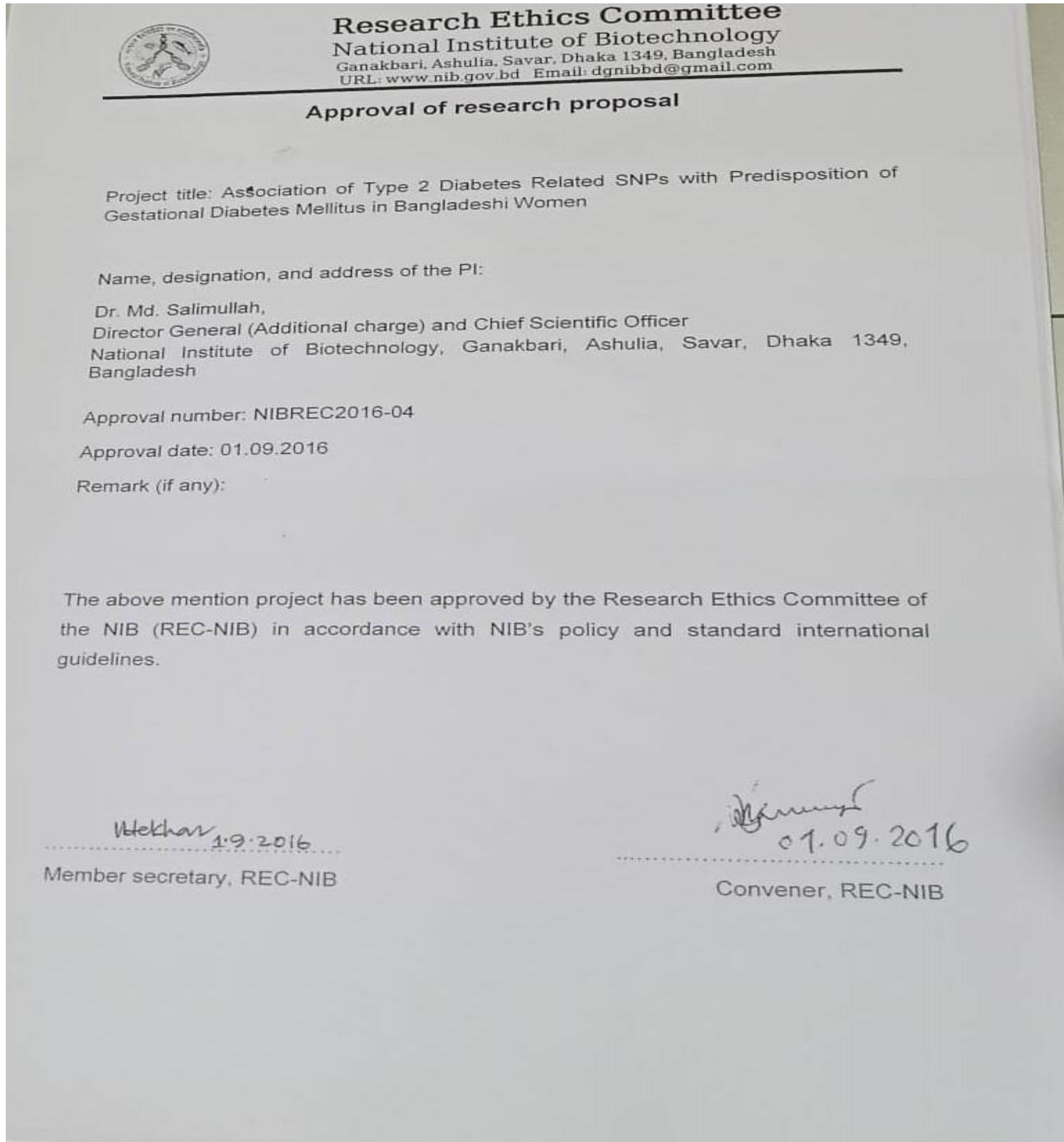
4. TBE buffer, used in gel electrophoresis (1X)

Components for preparing 1X TBE

Component	Amount
5× TBE	100 ml
ddH2O, volume to	500 ml

N.B. concentrated stock buffer was diluted just before use.

A5. Approval of research proposal



Research Ethics Committee
National Institute of Biotechnology
Ganakbari, Ashulia, Savar, Dhaka 1349, Bangladesh
URL: www.nib.gov.bd Email: dgnibbd@gmail.com

Approval of research proposal

Project title: Association of Type 2 Diabetes Related SNPs with Predisposition of Gestational Diabetes Mellitus in Bangladeshi Women

Name, designation, and address of the PI:

Dr. Md. Salimullah,
Director General (Additional charge) and Chief Scientific Officer
National Institute of Biotechnology, Ganakbari, Ashulia, Savar, Dhaka 1349, Bangladesh

Approval number: NIBREC2016-04

Approval date: 01.09.2016

Remark (if any):

The above mention project has been approved by the Research Ethics Committee of the NIB (REC-NIB) in accordance with NIB's policy and standard international guidelines.

Iftekhar 1.9.2016
Member secretary, REC-NIB

01.09.2016
Convener, REC-NIB

A6. Association of target SNPs with GDM

Table A1: Association of rs7756992 with GDM under different genetic models(n=468*)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant				**0.046**		0.061
A/A	124 (48.1)	78 (36.8)	1.00		1.00	
A/G	115 (44.6)	113(53.3)	**1.56 (1.06-2.29)**		**1.53 (1.03-2.27)**	
G/G	19 (7.4)	21 (9.9)	**1.76** (0.89-3.48)		1.81 (0.90-3.64)	
Dominant				**0.014**		**0.02**
A/A	124 (48.1)	78(36.8)	1.00		1.00	
A/G-G/G	134 (51.9)	134(63.2)	**1.59(1.10-2.30)**		**1.57 (1.07-2.29)**	
Recessive				0.33		0.28
A/A-A/G	239 (92.6)	191 (90.1)	1.00		1.00	
G/G	19 (7.4)	21 (9.9)	**1.38** (0.72-2.65)		1.44 (0.74-2.81)	
Overdominant				0.059		0.092
A/A-G/G	143 (55.4)	99(46.7)	1.00		1.00	
A/G	115 (44.6)	113(53.3)	**1.42** (0.99-2.04)		1.38 (0.95-2.01)	
Log-additive	---	---	**1.42 (1.06-1.90)**	**0.018**	**1.42 (1.05-1.91)**	**0.021**

[a] adjusted for gravidity and family history of diabetes

***Samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

Table A2: Association of rs7754840 with GDM under different genetic models(n=468*)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant				0.16		0.11
G/G	142 (55)	107 (50.5)	1.00		1.00	
G/C	103 (39.9)	85 (40.1)	1.10 (0.75-1.60)		1.16 (0.78-1.71)	
C/C	13 (5)	20 (9.4)	**2.04** (0.97-4.29)		**2.23 (1.04-4.75)**	
Dominant				0.32		0.2
G/G	142 (55)	107 (50.5)	1.00		1.00	
G/C-C/C	116 (45)	105 (49.5)	1.20 (0.83-1.73)		1.28 (0.88-1.86)	
Recessive				0.064		**0.048**
G/G-G/C	245 (95)	192 (90.6)	1.00		1.00	
C/C	13 (5)	20 (9.4)	**1.96** (0.95-4.05)		**2.09 (1.00-4.36)**	
Overdominant				0.97		0.8
G/G-C/C	155 (60.1)	127 (59.9)	1.00		1.05 (0.72-1.54)	
G/C	103 (39.9)	85 (40.1)	1.01 (0.70-1.46)		1.32 (0.98-1.79)	
Log-additive	---	---	1.26 (0.94-1.69)	0.12	1.05 (0.72-1.54)	0.066

[a] adjusted for gravidity and family history of diabetes

***Rest of the samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

Table A3: Association of rs8050136 with GDM under different genetic model(n=502*)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant C/C	143 (50.4%)	101 (46.3%)	1.00	0.28	1.00	0.37
A/C	126 (44.4%)	98 (45%)	1.10 (0.76-1.59)		1.09 (0.75-1.58)	
A/A	15 (5.3%)	19 (8.7%)	1.79 (0.87-3.70)		1.70 (0.81-3.55)	
Dominant C/C	143 (50.4%)	101 (46.3%)	1.00	0.37	1.00	0.44
A/C-A/A	141 (49.6%)	117 (53.7%)	1.17 (0.82-1.67)		1.15 (0.81-1.65)	
Recessive C/C-A/C	269 (94.7%)	199 (91.3%)	1.00	0.13	1.00	0.18
A/A	15 (5.3%)	19 (8.7%)	1.71 (0.85-3.45)		1.63 (0.80-3.33)	
Overdominant C/C-A/A	158 (55.6%)	120 (55%)	1.00	0.9	1.00	0.92
A/C	126 (44.4%)	98 (45%)	1.02 (0.72-1.46)		1.02 (0.71-1.46)	
Log-additive	---	---	1.22 (0.91-1.62)	0.18	1.19 (0.89-1.60)	0.24

[a] adjusted for gravidity and family history of diabetes

***3 Samples were excluded as genotypes could not be confirmed.**

Table A4: Association of *HSPAL1* gene rs2227956 with GDM under different genetic models(n=501*)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant T/T	210(74.2%)	162 (74.3%)	1.00	0.98	1.00	0.99
C/T	66 (23.3%)	50 (22.9%)	0.98 (0.64-1.50)		1.00 (0.65-1.54)	
C/C	7 (2.5%)	6 (2.8%)	1.11 (0.37-3.37)		1.08 (0.35-3.35)	
Dominant T/T	210 (74.2%)	162 (74.3%)	1.00	0.98	1.00	0.96
C/T-C/C	73 (25.8%)	56 (25.7%)	0.99 (0.66-1.49)		1.01 (0.67-1.52)	
Recessive T/T-C/T	276 (97.5%)	212 (97.2%)	1.00	0.85	1.00	0.89
C/C	7 (2.5%)	6 (2.8%)	1.12 (0.37-3.37)		1.08 (0.35-3.33)	
Overdominant T/T-C/C	217 (76.7%)	168 (77.1%)	1.00	0.92	1.00	1
C/T	66 (23.3%)	50 (22.9%)	0.98 (0.64-1.49)		1.00 (0.65-1.53)	
Log-additive	---	---	1.01 (0.71-1.43)	0.97	1.02 (0.71-1.45)	0.93

[a] adjusted for gravidity and family history of diabetes

***4 Samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

Table A5: Association of *PPARG* gene rs3856806 with GDM under different genetic models(n=502*)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant C/C	194 (68.1%)	148 (68.2%)	1.00	0.99	1.00	0.99
C/T	81 (28.4%)	61 (28.1%)	0.99 (0.66-1.47)		1.03 (0.69-1.54)	
T/T	10 (3.5%)	8 (3.7%)	1.05 (0.40-2.72)		1.01 (0.38-2.66)	
Dominant C/C	194 (68.1%)	148 (68.2%)	1.00	0.97	1.00	0.89
C/T-T/T	91 (31.9%)	69 (31.8%)	0.99 (0.68-1.45)		1.03 (0.70-1.51)	
Recessive C/C-C/T	275 (96.5%)	209 (96.3%)	1.00	0.92	1.00	1
T/T	10 (3.5%)	8 (3.7%)	1.05 (0.41-2.71)		1.00 (0.38-2.61)	
Overdominant C/C-T/T	204 (71.6%)	156 (71.9%)	1.00	0.94	1.00	0.89
C/T	81 (28.4%)	61 (28.1%)	0.98 (0.67-1.46)		1.03 (0.69-1.53)	
Log-additive	---	---	1.00 (0.73-1.38)	0.99	1.02 (0.74-1.41)	0.91

[a] adjusted for gravidity and family history of diabetes

***3 Samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

Table A6: Association of *PPARG* gene rs1801282 with GDM under different genetic models(n=505)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant **C/C**	213 (74.5%)	170 (77.6%)	1.00	0.7	1.00	0.69
C/G	68 (23.8%)	46 (21%)	0.85 (0.55-1.30)		0.90 (0.59-1.39)	
G/G	5 (1.8%)	3 (1.4%)	0.75 (0.18-3.19)		0.57 (0.13-2.47)	
Dominant C/C	213 (74.5%)	170 (77.6%)	1.00	0.41	1.00	0.54
C/G-G/G	73 (25.5%)	49 (22.4%)	0.84 (0.56-1.27)		0.88 (0.58-1.33)	
Recessive **C/C-C/G**	281 (98.2%)	216 (98.6%)	1.00	0.73	1.00	0.46
G/G	5 (1.8%)	3 (1.4%)	0.78 (0.18-3.30)		0.58 (0.13-2.51)	
Overdominant **C/C-G/G**	218 (76.2%)	173 (79%)	1.00	0.46	1.00	0.68
C/G	68 (23.8%)	46 (21%)	0.85 (0.56-1.30)		0.91 (0.59-1.41)	
Log-additive	---	---	0.85 (0.58-1.24)	0.4	0.86 (0.59-1.26)	0.45

[a] adjusted for gravidity and family history of diabetes

Table A7: Association of *TCF7L2* gene rs10885406 with GDM under different genetic models(n=505)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant						
A/A	111 (38.8%)	85(38.8%)	1.00	0.82	1.00	0.73
A/G	140 (49%)	111(50.7%)	1.04 (0.71-1.51)		1.01 (0.69-1.49)	
G/G	35 (12.2%)	23 (10.5%)	0.86 (0.47-1.56)		0.80 (0.43-1.48)	
Dominant						
A/A	111 (38.8%)	85 (38.8%)	1.00	1	1.00	0.87
A/G-G/G	175 (61.2%)	134 (61.2%)	1.00 (0.70-1.43)		0.97 (0.67-1.40)	
Recessive						
A/A-A/G	251 (87.8%)	196 (89.5%)	1.00	0.54	1.00	0.43
G/G	35 (12.2%)	23 (10.5%)	0.84 (0.48-1.47)		0.79 (0.45-1.41)	
Overdominant						
A/A-G/G	146 (51%)	108 (49.3%)	1.00	0.7	1.00	0.73
A/G	140 (49%)	111 (50.7%)	1.07(0.75-1.52)		1.06 (0.74-1.53)	
Log-additive	---	---	0.96(0.73-1.26)	0.77	0.93(0.71-1.23)	0.61

[a] adjusted for gravidity and family history of diabetes

Table A8: Association of *TCF7L2* gene rs7903146 variant with GDM under different genetic models(n=501*)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI) [a]	*P* value [a]
Codominant						
C/C	142(50%)	112(51.6%)	1.00	0.78	1.00	0.69
C/T	122(43%)	93(42.9%)	0.97 (0.67-1.39)		0.93 (0.64-1.35)	
T/T	20(7%)	12(5.5%)	0.76 (0.36-1.62)		0.72 (0.33-1.56)	
Dominant						
C/C	142(50%)	112(51.6%)	1.00	0.72	1.00	0.57
C/T-T/T	142(50%)	105(48.4%)	0.94 (0.66-1.34)		0.90 (0.63-1.29)	
Recessive						
C/C-C/T	264(93%)	205(94.5%)	1.00	0.49	1.00	0.44
T/T	20(7%)	12(5.5%)	0.77 (0.37-1.62)		0.75 (0.35-1.58)	
Overdominant						
C/C-T/T	162(57%)	124(57.1%)	1.00	0.98	1.00	0.85
C/T	122(43%)	93(42.9%)	1.00 (0.70-1.42)		0.96 (0.67-1.39)	
Log-additive	---	---	0.92 (0.69-1.23)	0.57	0.89 (0.66-1.20)	0.44

[a] adjusted for gravidity and family history of diabetes

***4 Samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

Table A9: Association of *TCF7L2* gene rs12255372 with GDM under different genetic models (n=502)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI)[a]	*P* value[a]
Codominant G/G	154 (54%)	95 (43.8%)	1.00	0.072	1.00	0.14
G/T	118(41.4%)	111 (51.1%)	**1.52 (1.06-2.19)**		1.45(1.00-2.10)	
T/T	13 (4.6%)	11 (5.1%)	1.37 (0.59-3.19)		1.39 (0.59-3.28)	
Dominant G/G	154 (54%)	95 (43.8%)	1.00	**0.023**	1.00	**0.046**
G/T-T/T	131 (46%)	122 (56.2%)	**1.51(1.06-2.15)**		**1.44 (1.01-2.07)**	
Recessive G/G-G/T	272 (95.4%)	206 (94.9%)	1.00	0.79	1.00	0.73
T/T	13 (4.6%)	11 (5.1%)	1.12 (0.49-2.54)		1.16 (0.50-2.69)	
Overdominant G/G-T/T	167 (58.6%)	106 (48.9%)	1.00	**0.03**	1.00	0.064
G/T	118 (41.4%)	111 (51.1%)	**1.48 (1.04-2.11)**		1.41 (0.98-2.02)	
Log-additive	---	---	**1.37 (1.01-1.85)**	**0.041**	1.33 (0.98-1.81)	0.067

[a] **adjusted for gravidity and family history of diabetes**

***3 Samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

Table A10: Association of *WFS1* gene rs10010131 with GDM under different genetic models(n=502)

Model	Control (%)	GDM (%)	OR (95% CI)	*P* value	OR (95% CI)[a]	*P* value[a]
Codominant **G/G**	153 (54.1%)	124 (56.6%)	1.00	0.32	1.00	0.25
A/G	113 (39.9%)	76(34.7%)	0.83 (0.57-1.21)		0.83 (0.57-1.21)	
A/A	17(6%)	19(8.7%)	1.38 (0.69-2.77)		1.50(0.74-3.04)	
Dominant **G/G**	153 (54.1%)	124 (56.6%)	1.00	0.57	1.00	0.62
A/G-A/A	130 (45.9%)	95 (43.4%)	0.90 (0.63-1.29)		0.91 (0.64-1.31)	
Recessive **G/G-A/G**	266 (94%)	200 (91.3%)	1.00	0.25	1.00	0.17
A/A	17 (6%)	19 (8.7%)	1.49 (0.75-2.93)		1.61 (0.81-3.22)	
Overdominant **G/G-A/A**	170 (60.1%)	143 (65.3%)	1.00	0.23	1.00	0.22
A/G	113 (39.9%)	76 (34.7%)	0.80 (0.55-1.15)		0.79 (0.54-1.15)	
Log-additive	---	---	1.00(0.76-1.33)	0.98	1.02(0.77-1.36)	0.87

[a] **adjusted for gravidity and family history of diabetes**

***3 Samples were excluded as heterozygosity and incomplete digestion could not be differentiated.**

A7. Cumulative Association of Target SNPs and FHO with GDM

Table A11: Cross classification interaction table of *CDKAL1* variants (rs7756992 and rs775 4840) and family history of T2DM

SNP	Models	Family history of T2DM (n=462*)						Interaction *P* value
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
rs7756992	Codominant A/A	83	49	1.00	41	29	1.20 (0.66-2.17)	0.1
	A/G	82	55	1.14 (0.70-1.86)	33	57	**2.93 (1.68-5.10)**	
	G/G	14	9	1.09 (0.44-2.70)	5	12	**4.07 (1.35-12.23)**	
	Dominant A/A	83	49	1.00	41	29	1.20 (0.66-2.17)	**0.038**
	A/G-G/G	96	64	1.13 (0.70-1.81)	38	69	**3.08 (1.81-5.23)**	
	Recessive A/A-A/G	162	102	1.00	73	85	**1.89 (1.27-2.83)**	**0.32**
	G/G	14	9	1.08 (0.45-2.61)	5	12	**4.11 (1.40-12.12)**	
	Overdominant A/A-G/G	95	57	1.00	46	40	1.48 (0.86-2.53)	**0.12**
	A/G	81	54	1.09 (0.68-1.76)	32	57	**2.99 (1.73-5.16)**	
rs7754840	Codominant G/G	95	57	1.00	47	49	**1.74 (1.04-2.92)**	0.53
	G/C	73	45	1.03 (0.63-1.69)	30	40	**2.22 (1.25-3.95)**	
	C/C	11	11	1.67 (0.68-4.09)	2	9	**7.50 (1.57-35.94)**	
	Dominant G/G	95	57	1.00	47	49	**1.74 (1.04-2.92)**	0.47
	G/C-C/C	84	56	1.11 (0.69-1.78)	32	49	**2.55 (1.47-4.44)**	
	Recessive G/G-G/C	165	100	1.00	76	88	**1.95 (1.31-2.90)**	**0.31**
	C/C	11	11	1.63 (0.68-3.91)	2	9	**7.74 (1.63-36.76)**	
	Overdominant	105	66	1.00	48	57	**1.91 (1.16-3.13)**	**0.72**
		71	45	1.00 (0.61-1.62)	30	40	**2.19 (1.24-3.87)**	

[a] adjusted for gravidity

***Samples were excluded due to missing information of family history of diabetes.**

Table A12: Cross classification interaction table of *FTO* variant rs8050136 and family history of T2DM under different genetic models

Models	Genotypes	Family history of T2DM (n=502*)						Interaction *P* value [a]
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	C/C	102	54	1.00	41	47	**2.22 (1.30-3.80)**	0.55
	A/C	82	54	1.25 (0.77-2.01)	44	44	**1.95 (1.14-3.34)**	
	A/A	8	9	2.26 (0.82-6.24)	7	10	2.67 (0.96-7.44)	
Dominant	C/C	102	54	1.00	41	47	**2.22 (1.30-3.79)**	0.33
	A/C-A/A	90	63	1.33 (0.84-2.12)	51	54	**2.05 (1.23-3.41)**	
Recessive	C/C-A/C	184	108	1.00	85	91	**1.88 (1.28-2.75)**	0.52
	A/A	8	9	2.04 (0.76-5.47)	7	10	2.40 (0.88-6.53)	
Overdominant	C/C-A/A	110	63	1.00	48	57	**2.10 (1.28-3.45)**	0.43
	A/C	82	54	1.15 (0.72-1.82)	44	44	**1.79 (1.06-3.03)**	

[a] adjusted for gravidity
***Samples were excluded due to missing information of family history of diabetes.**

Table A13: Cross classification interaction table of *HSPAL1* gene variant rs2227956 and family history of T2DM

Models	Genotypes	Family history of T2DM (n=495*)						Interaction *P* value [a]
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	T/T	144	82	1.00	66	80	2.17 (1.41-3.32)	0.21
	C/T	44	30	1.20 (0.70-2.06)	22	20	1.63 (0.84-3.18)	
	C/C	3	4	2.38 (0.51-10.98)	4	2	0.91 (0.16-5.12)	
Dominant	T/T	144	82	1.00	66	80	2.17 (1.41-3.32)	0.16
	C/T-C/C	47	34	1.28 (0.76-2.15)	26	22	1.52 (0.81-2.86)	

Recessive	T/T-C/T	188	112	1.00	88	100	1.94 (1.34-2.82)	0.16
	C/C	3	4	2.27 (0.49-10.42)	4	2	0.87 (0.16-4.86)	
Overdominant	T/T-C/C	147	86	1.00	70	82	2.04 (1.34-3.10)	0.36
	C/T	44	30	1.17 (0.68-2.00)	22	20	1.59 (0.82-3.09)	

[a] adjusted for gravidity

***Samples were excluded due to missing information of family history of diabetes.**

Table A14: Cross classification interaction table of *PPARG* gene rs3856806 variant and family history of T2DM

Models	**Genotypes**	**Family history of T2DM (n=502*)**						**Interaction *P* value [a]**
		No			**Yes**			
		Control	**GDM**	**OR (95% CI)**	**Control**	**GDM**	**OR (95% CI)**	
Codominant	C/C	131	75	1.00	63	73	**2.04 (1.31-3.18)**	0.85
	C/T	56	36	1.12 (0.68-1.87)	25	25	1.82 (0.97-3.41)	
	T/T	6	4	1.11 (0.30-4.09)	4	4	1.82 (0.44-7.56)	
Dominant	C/C	131	75	1.00	63	73	**2.04 (1.31-3.18)**	0.57
	C/T-T/T	62	40	1.12 (0.69-1.83)	29	29	**1.82 (1.01-3.30)**	
Recessive	C/C-C/T	187	111	1.00	88	98	**1.91 (1.31-2.78)**	0.88
	T/T	6	4	1.07 (0.29-3.91)	4	4	1.76 (0.43-7.23)	
Overdominant	C/C-T/T	137	79	1.00	67	77	**2.02 (1.31-3.11)**	0.6
	C/T	56	36	1.12 (0.68-1.85)	25	25	1.81 (0.97-3.39)	

[a] adjusted for gravidity

***Samples were excluded due to missing information of family history of diabetes.**

Table A15: Cross classification interaction table of *PPARG* gene rs1801282 variant and fam ily history of T2DM

Models	Genotypes	Family history of T2DM (n=505)						Interaction *P* value [a]
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	C/C	141	89	1.00	72	81	1.80 (1.19-2.73)	0.87
	C/G	51	27	0.83 (0.49-1.43)	17	19	1.89 (0.93-3.86)	
	G/G	2	1	0.66 (0.06-7.42)	3	2	0.96 (0.16-5.90)	
Dominant	C/C	141	89	1.00	72	81	1.80 (1.18-2.72)	0.72
	C/G-G/G	53	28	0.83 (0.49-1.41)	20	21	1.74 (0.89-3.41)	
Recessive	C/C-C/G	192	116	1.00	89	100	1.90 (1.31-2.75)	0.86
	G/G	2	1	0.69 (0.06-7.74)	3	2	1.00 (0.16-6.15)	
Overdominant	C/C-G/G	143	90	1.00	75	83	1.77 (1.17-2.67)	0.59
	C/G	51	27	0.84 (0.49-1.44)	17	19	1.90 (0.93-3.88)	

[a] adjusted by gravidity
***Samples were excluded due to missing information of family history of diabetes.**

Table A16: Cross classification interaction table of *TCF7L2* gene rs10885406 variant and fa mily history of T2DM

Models	Genotypes	Family history of T2DM (n=505)						Interaction *P* value [a]
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	A/A	85	43	1.00	26	42	**3.42 (1.84-6.34)**	**0.053**
	A/G	87	61	1.48 (0.90-2.44)	53	50	**1.95 (1.14-3.35)**	

	G/G	22	13	1.20 (0.55-2.63)	13	10	1.63 (0.66-4.05)	
Dominant	A/A	85	43	1.00	26	42	**3.42 (1.84-6.34)**	**0.015**
	A/G-G/G	109	74	1.42 (0.88-2.29)	66	60	**1.89 (1.13-3.15)**	
Recessive	A/A-A/G	172	104	1.00	79	92	**1.96 (1.33-2.89)**	0.53
	G/G	22	13	0.97 (0.47-2.02)	13	10	1.31 (0.55-3.11)	
Overdominant	A/A-G/G	107	56	1.00	39	52	**2.71 (1.59-4.62)**	**0.055**
	A/G	87	61	1.42 (0.89-2.27)	53	50	**1.88 (1.13-3.12)**	

[a] adjusted for gravidity

Table A17: Cross classification interaction table of *TCF7L2* gene rs7903146 variant and family history of T2DM

Models	**Genotypes**	**Family history of T2DM (n=501*)**						**Interaction *P* value [a]**
		No			**Yes**			
		Control	**GDM**	**OR (95% CI)**	**Control**	**GDM**	**OR (95% CI)**	
Codominant	**C/C**	103	64	1.00	39	48	**2.05 (1.21-3.48)**	0.87
	C/T	77	47	1.01 (0.62-1.63)	45	46	**1.69 (1.01-2.85)**	
	T/T	12	5	0.71 (0.24-2.12)	8	7	1.46 (0.50-4.25)	
Dominant	**C/C**	103	64	1.00	39	48	**2.05 (1.21-3.48)**	0.63
	C/T-T/T	89	52	0.97 (0.61-1.54)	53	53	**1.66 (1.01-2.72)**	
Recessive	**C/C-C/T**	180	111	1.00	84	94	**1.85 (1.27-2.71)**	0.89
	T/T	12	5	0.71 (0.24-2.07)	8	7	1.46 (0.51-4.16)	
Overdominant	**C/C-T/T**	115	69	1.00	47	55	**2.01 (1.23-3.29)**	0.64
	C/T	77	47	1.04 (0.65-1.66)	45	46	**1.74 (1.05-2.91)**	

[a] adjusted for gravidity

***Samples were excluded due to missing information of family history of diabetes.**

Table A18: Cross classification interaction table of *TCF7L2* gene variant rs12255372 and family history of T2DM

Models	Genotypes	Family history of T2DM (n=502*)						Interaction P value [a]
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	G/G	109	56	1.00	45	39	1.72 (1.00-2.94)	0.71
	G/T	74	56	1.45 (0.90-2.34)	44	55	**2.49 (1.49-4.16)**	
	T/T	10	5	1.05 (0.34-3.23)	3	6	3.80 (0.91-15.86)	
Dominant	G/G	109	56	1.00	45	39	1.72 (1.00-2.94)	0.87
	G/T-T/T	84	61	1.41 (0.89-2.24)	47	61	**2.57 (1.56-4.25)**	
Recessive	G/G-G/T	183	112	1.00	89	94	**1.77 (1.22-2.58)**	0.43
	T/T	10	5	0.88 (0.29-2.67)	3	6	3.21 (0.78-13.16)	
Overdominant	G/G-T/T	119	61	1.00	48	45	**1.84 (1.10-3.08)**	0.84
	G/T	74	56	1.45 (0.91-2.31)	44	55	**2.48 (1.49-4.11)**	

[a] adjusted for gravidity
***Samples were excluded due to missing information of family history of diabetes.**

Table A19: Cross classification interaction table of *WFS1* gene variant rs10010131 and family history of T2DM

Models	Genotypes	Family history of T2DM (n=502*)						Interaction P value [a]
		No			Yes			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	G/G	105	63	1.00	48	61	**2.28 (1.39-3.76)**	0.4
	A/G	74	40	0.96	39	36	1.53 (0.88-2.67)	

				(0.58-1.58)				
	A/A	12	14	1.98 (0.86-4.57)	5	5	1.79 (0.49-6.48)	
Dominant	**G/G**	105	63	1.00	48	61	**2.29 (1.39-3.76)**	0.21
	A/G-A/A	86	54	1.11 (0.69-1.76)	44	41	1.56 (0.92-2.66)	
Recessive	**G/G-A/G**	179	103	1.00	87	97	**1.97 (1.35-2.88)**	0.31
	A/A	12	14	2.01 (0.89-4.54)	5	5	1.81 (0.51-6.46)	
Overdominant	**G/G-A/A**	117	77	1.00	53	66	**2.03 (1.27-3.25)**	0.54
	A/G	74	40	0.87 (0.54-1.41)	39	36	1.40 (0.81-2.39)	

[a] adjusted for gravidity
***Samples were excluded due to missing information of family history of diabetes.**

A8. Cumulative association of target SNP and Gravidity with GDM

Table A20: Cross classification interaction table of *CDKAL1* variants (rs7756992 and rs7754840) and gravidity

SNP	**Models**	**Gravidity(n=465*)**						**Interaction *P* value**
		Primigravida			**Multigravida**			
		Control	**GDM**	**OR (95% CI)**	**Control**	**GDM**	**OR (95% CI)**	
rs7756992	Codominant A/A A/G G/G	56	29	1.00	67	47	1.35 (0.76-2.43)	0.6
		53	36	1.31 (0.71-2.43)	61	76	**2.41 (1.37-4.22)**	
		10	11	2.12 (0.81-5.58)	9	10	2.15 (0.78-5.87)	
	Dominant A/A A/G-G/G	56	29	1.00	67	47	1.35 (0.76-2.43)	0.62
		63	47	1.44 (0.80-2.59)	70	86	**2.37 (1.37-4.10)**	
		107	65	1.00	128	122	**1.63**	

	Recessive						**(1.09-2.44)**	**0.48**
	A/A-A/G G/G	10	11	1.80 (0.71-4.53)	9	10	1.83 (0.70-4.80)	
	Overdominant A/A-G/G	65	40	1.00	76	57	1.26 (0.74-2.14)	**0.32**
	A/G	52	36	1.10 (0.61-1.98)	61	75	**2.04 (1.20-3.45)**	
rs7754840	Codominant	64	39	1.00	77	65	1.39 (0.83-2.32)	
		49	30	1.00 (0.55-1.84)	53	55	1.70 (0.98-2.95)	
		6	7	1.91 (0.60-6.11)	7	13	**3.05 (1.12-8.30)**	0.88
	Dominant	64	39	1.00	77	65	1.39 (0.83-2.32)	
		55	37	1.10 (0.62-1.96)	60	68	**1.86 (1.10-3.15)**	0.61
	Recessive G/G-G/C	111	69	1.00	130	119	**1.52 (1.02-2.27)**	**0.88**
	C/C	6	7	1.94 (0.61-6.10)	7	13	**3.31 (1.24-8.83)**	
	Overdominant	69	46	1.00	84	77	1.39 (0.85-2.28)	0.53
		48	30	0.91 (0.50-1.65)	53	55	1.62 (0.95-2.78)	

***Samples were excluded due to missing information of gravidity.**

Table A21: Cross classification interaction table of *FTO* variant rs8050136 and gravidity under different genetic models

Models	Genotypes	Gravidity(n=502*)						Interaction *P* value [a]
		Primigravida			Multigravida			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	C/C	57	45	1.00	86	56	0.84 (0.50-1.42)	**0.0068**
	C/A	67	29	**0.53 (0.29-0.95)**	59	69	1.51 (0.89-2.56)	
	A/A	8	7	1.10 (0.37-3.30)	7	12	1.96 (0.70-5.44)	
Dominant	C/C	57	45	1.00	86	56	0.84 (0.50-1.42)	**0.0021**
	C/A-A/A	75	36	0.59	66	81	1.56	

				(0.33-1.03)			(0.93-2.60)	
Recessive	C/C-A/C	124	74	1.00	145	125	**1.50 (1.03-2.19)**	0.81
	A/A	8	7	1.48 (0.51-4.30)	7	12	2.64 (0.99-7.09)	
Overdominant	C/C-A/A	65	52	1.00	93	68	0.92 (0.57-1.49)	**0.0025**
	A/C	67	29	**0.52 (0.29-0.92)**	59	69	1.49 (0.90-2.48)	

[a] adjusted for family history of diabetes

***Samples were excluded due to missing information of gravidity.**

Table A22: Cross classification interaction table of *HSPAL1* gene variant rs2227956 and gr avidity

Models	Genotypes	Gravidity(n=495*)						Interaction *P* value [a]
		Primigravida			Multigravida			
		Contr ol	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	**T/T**	99	59	1.00	111	103	**1.59 (1.04-2.44)**	0.82
	C/T	31	19	1.05 (0.54-2.03)	35	31	1.55 (0.86-2.80)	
	C/C	3	3	1.60 (0.31-8.35)	4	3	1.25 (0.26-5.87)	
Dominant	**T/T**	99	59	1.00	111	103	**1.59 (1.04-2.44)**	**0.74**
	C/T-C/C	34	22	1.10 (0.58-2.06)	39	34	1.52 (0.86-2.69)	
Recessive	**T/T-C/T**	130	78	1.00	146	134	**1.57 (1.08-2.27)**	**0.54**
	C/C	3	3	1.58 (0.31-8.19)	4	3	1.23 (0.26-5.76)	
Overdominant	**T/T-C/C**	102	62	1.00	115	106	**1.55 (1.02-2.36)**	**0.92**
	C/T	31	19	1.03 (0.53-1.99)	35	31	1.53 (0.85-2.74)	

[a] adjusted for family history of diabetes

***Samples were excluded due to missing information of gravidity.**

Table A23: Cross classification interaction table of *PPARG* gene rs3856806 variant and gra vidity

Models	Genotypes	Gravidity (n=502*)						Interaction P value [a]
		Primigravida			Multigravida			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	C/C	88	54	1.00	106	94	1.46 (0.94-2.28)	0.91
	C/T	40	23	0.94 (0.50-1.75)	41	38	1.61 (0.92-2.84)	
	T/T	4	3	1.10 (0.23-5.20)	6	5	1.40 (0.40-4.89)	
Dominant	C/C	88	54	1.00	106	94	1.46 (0.94-2.28)	0.75
	C/T-T/T	44	26	0.95 (0.52-1.74)	47	43	1.59 (0.92-2.73)	
Recessive	C/C-C/T	128	77	1.00	147	132	1.53 (1.06-2.23)	0.85
	T/T	4	3	1.12 (0.24-5.24)	6	5	1.43 (0.42-4.92)	
Overdominant	C/C-T/T	92	57	1.00	112	99	1.45 (0.94-2.24)	0.69
	C/T	40	23	0.93 (0.50-1.73)	41	38	1.61 (0.92-2.81)	

[a] adjusted for family history of diabetes

***Samples were excluded due to missing information of gravidity.**

Table A24: Cross classification interaction table of *PPARG* gene rs1801282 variant and gra vidity

Models	**Genotypes**	**Gravidity (n=505)**						**Interaction *P* value [a]**
		Primigravida			**Multigravida**			
		Control	**GDM**	**OR (95% CI)**	**Control**	**GDM**	**OR (95% CI)**	
Codominant	C/C	98	64	1.00	115	106	1.43 (0.94-2.16)	0.16
	C/G	36	16	0.68 (0.35-1.34)	32	30	1.59 (0.87-2.89)	
	G/G	0	1	---	5	2	0.54 (0.10-2.94)	
Dominant	C/C	98	64	1.00	115	106	1.43 (0.94-2.16)	0.45

	C/G-G/G	36	17	0.72 (0.37-1.40)	37	32	1.43 (0.80-2.55)	
Recessive	C/C-C/G	134	80	1.00	147	136	1.60 (1.11-2.30)	0.12
	G/G	0	1	---	5	2	0.59 (0.11-3.19)	
Overdominant	C/C-G/G	98	65	1.00	120	108	1.37 (0.91-2.07)	0.24
	C/G	36	16	0.67 (0.34-1.32)	32	30	1.57 (0.86-2.85)	

[a] adjusted by family history of diabetes

Table A25: Cross classification interaction table of *TCF7L2* gene rs10885406 variant and gr avidity

Models	Genotypes	Gravidity (n=505)						Interaction *P* value [a]
		Primigravida			Multigravida			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	**A/A**	46	29	1.00	65	56	1.48 (0.81-2.68)	0.62
	A/G	70	45	1.04 (0.57-1.90)	70	66	1.49 (0.83-2.66)	
	G/G	18	7	0.60 (0.22-1.63)	17	16	1.56 (0.67-3.59)	
Dominant	**A/A**	46	29	1.00	65	56	1.47 (0.81-2.67)	0.85
	A/G-G/G	88	52	0.95 (0.53-1.70)	87	82	1.50 (0.86-2.63)	
Recessive	**A/A-A/G**	116	74	1.00	135	122	1.45 (0.98-2.13)	
	G/G	18	7	0.59 (0.23-1.48)	17	16	1.52 (0.72-3.22)	0.33
Overdominant	**A/A-G/G**	64	36	1.00	82	72	1.68 (1.00-2.85)	0.66
	A/G	70	45	1.17 (0.67-2.06)	70	66	1.68 (0.98-2.87)	

[a] adjusted for family history of diabetes

Table A26: Cross classification interaction table of *TCF7L2* gene rs7903146 variant and gra vidity

		Gravidity (n=501*)	

Models	Genotypes	Primigravida			Multigravida			Interaction P value [a]
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	C/C	62	41	1.00	80	71	1.40 (0.84-2.35)	0.65
	C/T	59	36	0.90 (0.50-1.60)	63	57	1.34 (0.78-2.29)	
	T/T	12	4	0.48 (0.14-1.60)	8	8	1.41 (0.48-4.12)	
Dominant	C/C	62	41	1.00	80	71	1.40 (0.84-2.35)	0.69
	C/T-T/T	71	40	0.82 (0.47-1.44)	71	65	1.35 (0.80-2.28)	
Recessive	C/C-C/T	121	77	1.00	143	128	1.45 (0.99-2.11)	
	T/T	12	4	0.50 (0.16-1.64)	8	8	1.49 (0.53-4.19)	0.36
Overdominant	C/C-T/T	74	45	1.00	88	79	1.54 (0.95-2.50)	
	C/T	59	36	0.98 (0.56-1.72)	63	57	1.47 (0.87-2.47)	0.94

[a] adjusted for family history of diabetes
***Samples were excluded due to missing information of gravidity.**

Table A27: Cross classification interaction table of *TCF7L2* gene variant rs12255372 and gr avidity

Models	Genotypes	Gravidity (n=502*)						Interaction P value [a]
		Primigravida			Multigravida			
		Control	GDM	OR (95% CI)	Control	GDM	OR (95% CI)	
Codominant	G/G	69	39	1.00	85	56	1.19 (0.70-2.00)	0.4
	G/T	57	38	1.10 (0.62-1.96)	61	73	**2.08 (1.23-3.52)**	
	T/T	8	4	0.94 (0.26-3.35)	5	7	2.31 (0.68-7.88)	
Dominant	G/G	69	39	1.00	85	56	1.19 (0.70-2.00)	0.19
	G/T-T/T	65	42	1.08 (0.62-1.89)	66	80	**2.10 (1.26-3.52)**	

Models	Genotypes	Primigravida Control	Primigravida GDM	Primigravida OR (95% CI)	Multigravida Control	Multigravida GDM	Multigravida OR (95% CI)	Interaction P value [a]
Recessive	**G/G-G/T**	126	77	1.00	146	129	**1.50 (1.03-2.18)**	0.57
	T/T	8	4	0.90 (0.26-3.12)	5	7	2.21 (0.67-7.29)	
Overdominant	**G/G-T/T**	77	43	1.00	90	63	1.26 (0.77-2.07)	0.28
	G/T	57	38	1.11 (0.63-1.94)	61	73	**2.10 (1.26-3.49)**	

[a] adjusted for family history of diabetes

***Samples were excluded due to missing information of gravidity.**

Table A28: Cross classification interaction table of *WFS1* gene rs10010131 variant and grav idity

Models	**Genotypes**	**Gravidity (n=502*)**						**Interaction *P* value [a]**
		Primigravida			**Multigravida**			
		Control	**GDM**	**OR (95% CI)**	**Control**	**GD M**	**OR (95% CI)**	
Codominant	**G/G**	72	44	1.00	81	80	**1.80 (1.09-2.95)**	0.64
	A/G	53	30	1.03 (0.57-1.87)	60	46	1.27 (0.74-2.19)	
	A/A	8	7	1.59 (0.53-4.76)	9	12	2.60 (1.00-6.78)	
Dominant	**G/G**	72	44	1.00	81	80	**1.79 (1.09-2.94)**	0.4
	A/G-A/A	61	37	1.10 (0.63-1.93)	69	58	1.43 (0.85-2.40)	
Recessive	**G/G-A/G**	125	74	1.00	141	126	**1.55 (1.06-2.26)**	0.94
	A/A	8	7	1.56 (0.54-4.55)	9	12	**2.55 (1.01-6.42)**	
Overdominant	**G/G-A/A**	80	51	1.00	90	92	**1.77 (1.11-2.82)**	0.36
	A/G	53	30	0.97 (0.55-1.74)	60	46	1.20 (0.71-2.04)	

[a] adjusted for family history of diabetes
***Samples were excluded due to missing information of gravidity.**

A9.

Table A29: Comparisons of variables between trimester groups within cases, controls and all participants

Variables	Groups	1st Trimester (0-13 weeks)	2nd Trimester (14-26 weeks)	3rd Trimester (27 weeks and above)	*P* value
Age	Control	26.27±5.52	25.21±4.45	25.55±4.82	0.517
	GDM	27.51±4.86	27.24±4.70	27.79±4.56	0.746
	All	26.94±5.17	25.98±4.64	26.50±4.83	0.299
BMI	**Control**	**23.6±3.03**	**25.09±3.59**	**25.71±4.13**	**0.019**
	GDM	25.62±5.26	26.38±3.70	27.29±3.84	0.096
	All	24.63±4.40	25.58±3.67	26.38±4.08	0.004
SBP	Control	106.82±11.44	108.41±12.61	109.39±11.69	0.522
	GDM	108.78±12.66	110.42±12.70	108.64±11.13	0.608
	All	107.86±12.06	109.17±12.65	109.07±11.43	0.718
DBP	**Control**	**68.03±8.19**	**67.37±8.91**	**70.46±9.02**	**0.021**
	GDM	72.30±9.69	70.14±9.14	70.30±8.86	0.457
	All	70.29±9.20	68.42±9.07	70.39±8.93	0.070
FBS	Control	4.32±0.45	4.38±0.46	4.27±0.47	0.211
	GDM	5.28±0.64	5.11±0.69	5.12±0.71	0.383
	All	4.84±0.74	4.66±0.66	4.64±0.72	0.104
OBS	Control	7.63±6.66	7.47±6.32	7.51±6.51	0.806
	GDM	9.52±1.70	9.97±1.70	9.95±1.41	0.294
	All	8.65±1.76	8.43±1.88	8.56±1.77	0.628
TBS	Control	6.66±1.04	6.32±1.06	6.51±1.01	0.167
	GDM	7.83±1.55	8.33±1.48	8.42±1.45	0.102
	All	7.31±1.46	7.08±1.58	7.32±1.54	0.246

Table A30: Comparisons of variables between control and GDM groups within each trimester group

Variables	Trimester	Control	GDM	*P* Value
Age	1st	26.27±5.52	27.51±4.86	0.314
	2nd	25.21±4.45	27.24±4.70	**0.003**
	3rd	25.55±4.82	27.79±4.56	**0.000**
BMI	1st	23.6±3.03	25.62±5.26	0.064
	2nd	25.09±3.59	26.38±3.70	**0.022**
	3rd	25.71±4.13	27.29±3.84	**0.004**
SBP	1st	106.82±11.44	108.78±12.66	0.500
	2nd	108.41±12.61	110.42±12.70	0.291

	3rd	109.39±11.69	108.64±11.13	0.622
DBP	1st	68.03±8.19	72.30±9.69	**0.052**
	2nd	67.37±8.91	70.14±9.14	**0.042**
	3rd	70.46±9.02	70.30±8.86	0.897
FBS	1st	4.32±0.45	5.28±0.64	**0.000**
	2nd	4.38±0.46	5.11±0.69	**0.000**
	3rd	4.27±0.47	5.12±0.71	**0.000**
OBS	1st	7.63±6.66	9.52±1.70	**0.000**
	2nd	7.47±6.32	9.97±1.70	**0.000**
	3rd	7.51±6.51	9.95±1.41	**0.000**
TBS	1st	6.66±1.04	7.83±1.55	**0.001**
	2nd	6.32±1.06	8.33±1.48	**0.000**
	3rd	6.51±1.01	8.42±1.45	**0.000**