Biochemical Genetics (2019) 57:20–33
https://doi.org/10.1007/s10528-018-9872-7

ORIGINAL ARTICLE

Functional Polymorphism Located in the microRNA Binding Site of the Insulin Receptor (*INSR*) Gene Confers Risk for Type 2 Diabetes Mellitus in the Bangladeshi Population

Mahrima Parvin[1] · Farhana Jahan[1] · Pankaj Kumar Sarkar[2] · Zakir Hossain Howlader[3] · A. H. M. Nurun Nabi[1] · Md. Ismail Hosen[1,3]

Received: 12 March 2018 / Accepted: 25 June 2018 / Published online: 4 July 2018

Abstract

Bangladesh has the second largest number of adults with diabetes in South Asia. Compelling evidence suggest that miRNAs contribute to the etiology of Type 2 diabetes mellitus (T2DM) by regulating many aspects of glucose homeostasis. Hence, we hypothesized that genetic polymorphisms in the diabetes-related miRNA target-binding sites could be associated with the risk of T2DM in Bangladesh. The reference Single nucleotide polymorphism (SNP) data from the Insulin Receptor (*INSR*) gene were downloaded from the ENSEMBL genome browser release 88 and further analyzed in silico for identifying SNPs with deleterious effect and clinical relationships. Further, case–control study using the microRNA-binding site polymorphism rs1366600 (T > C) located at the 3′ *UTR* of the *INSR* gene was carried out in 217 T2DM patients and 237 healthy controls from Bangladesh. Genotyping was performed using the real time PCR based allele discrimination method. The results showed that the minor allele 'C' is associated with increased risk of T2DM [Odds ratio (OR) 1.87; 95% confidence intervals (CI) 1.28–2.74; $P = 0.0010$]. When we dissected our analysis to include the dominant model (CC + TC genotype against the TT genotype), we found that the CC and TC genotypes were associated with increased risk of T2DM in Bangladeshi population (OR 2.01; 95% CI 1.31–3.07; $P = 0.0012$). However, in recessive model (CC vs TT + TC); the effect was not statistically significant (OR 2.23; 95% CI 0.66–7.51; $P = 0.1848$). Stratification of our data based on the gender of the cases and controls showed similar degree of risk association with respect to different genotypes and alleles. Our study showed that the miRNA binding site polymorphism rs1366600 located at the 3′-UTR region of the *INSR* gene is associated with increased risk of T2DM in Bangladeshi individuals.

Keywords miRNA · Allelic discrimination assay · In silico analysis · INSR

The original version of this article was revised: The co-author name should be Farhana Jahan instead of Farhan Jahan. This has been corrected in this version.

✉ Md. Ismail Hosen
ismail.hosen@du.ac.bd

Extended author information available on the last page of the article

Introduction

Type 2 diabetes mellitus (T2DM) is a progressive metabolic disease characterized by insulin resistance and/or functional loss of pancreatic beta cells leading to hyperglycemic condition (Kahn et al. 2014). The burden of diabetes and its complications are rising in inflated fashion over the world in the past two decades and is expected to affect more than 500 million adults by 2030 (Whiting et al. 2011; Pradeepa and Mohan 2017). About 80% of the adults with diabetes live in low- and middle-income countries and the prevalence of diabetes is estimated to increase by 71% in this region by 2035 (Guariguata et al. 2014). The prevalence of diabetes in Bangladesh is high and expected to increase up to 50% by next 15 years and will be the 8th highest diabetic populous country in the world (Akter et al. 2014; Guariguata et al. 2014; Chowdhury et al. 2015).

The most common risk factors for T2DM include the interaction among genetic, environmental, and other risk factors of obesity and hypertension (Wu et al. 2014). Genetic predisposition studies are providing with novel insights on the mechanisms of T2DM (De Silva and Frayling 2010; Grarup et al. 2014). The results of these studies have significantly improved our knowledge on the pathogenesis and provided promising candidates for use in diagnostics and prognostics approaches (Lyssenko and Groop 2009).

microRNAs (miRNAs) are small RNA molecules that, by regulating gene expression, modulate diverse biological activities related to T2DM (carbohydrate and lipid metabolisms, insulin secretion and the adipocytokine signaling pathway) (Zhou and Yang 2012; He et al. 2017; Wu and Miller 2017). Genetic polymorphisms in the miRNA binding sites have been shown to influence the susceptibility to T2DM in different population (Gong et al. 2014; Goda et al. 2015a; Moszyńska et al. 2017). However, there is no such study in Bangladeshi population to portrait the association of miRNA binding site polymorphisms with T2DM.

The current study aims at underscoring the impact of a miRNA binding site polymorphism in the insulin receptor (*INSR*) gene with the susceptibility of T2DM in Bangladeshi individuals. To accomplish this, first we carried out an in silico analysis to identify SNPs that are causing the most damage to the structural features of the INSR protein, thus negatively affecting its biological functions. Then a selected miRNA binding site polymorphism (rs1366600) is genotyped in Bangladeshi T2DM patients and healthy controls via real time PCR based allelic discrimination assay.

Materials and Methods

Retrieval of SNP Data from the *INSR* Gene

The detailed information about the SNPs (including the allelic frequency; types of SNPs etc.) present in the *INSR* gene was retrieved from the Ensembl genome

browser version 88 (Aken et al. 2016). In Ensembl genome browser version 88 (released on March 2017), we used the search term INSR in Homo sapiens. In the gene information page, we selected the variant table track and the data for the variants were downloaded. From the downloaded SNP dataset for INSR, missense, 5′ UTR SNP, and 3′ UTR SNP were sorted and their corresponding information was collected.

Analysis of Functional Significance of SNPs in the Untranslated Region of the *INSR* Gene

RegulomeDB

RegulomeDB is a huge collection of data about regulatory variants of human genome based on high-throughput experimental data from ENCODE and other sources to identify putative regulatory potential and identify functional variants. These data sources are combined into a powerful tool that scores variants to help separate functional variants from a large pool and provides a small set of putative sites with testable hypotheses as to their function (Boyle et al. 2012). The input for regulomeDB was 3′ UTR SNP IDs. The results showed regulomeDB score which refers to the available datatypes for a single genomic coordinate on eQTL, TF binding, matched TF motif, matched DNase Footprint, and DNase peak. The supporting data list all the DNA features and regulatory regions (TF binding sites, DNase Footprinting, DNase sensitivity, Chromatin States, eQTLs, Differentially methylated regions, Manually curated regions, Validated functional SNPs) that have been identified to contain the input coordinate (Boyle et al. 2012).

PolymiRTS

The PolymiRTS v 3.0 (Polymorphisms in miRNAs and miRNA *t*arget *s*ites) is a complete database was developed to systematically identify DNA polymorphisms in miRNAs and miRNA target sites and elucidate their potential links to molecular, physiological, behavioral, and disease phenotypes (Bhattacharya et al. 2014). Gene symbol was search information and in the result a list of 3′ UTR SNPs of the gene creating or destructing miRNA site was obtained with miRNA IDs. The SNPs were further selected on the basis of 0.005–0.5 global minor allele frequency.

Study Population, Sample Collection and Genomic DNA Extraction

The population of this case control study consisted of 217 T2DM patients and 237 healthy controls of Bangladeshi origin. These diabetic patients were recruited from Dinajpur Diabetes O Swasthoseba Hospital, Bangladesh. The healthy controls were randomly selected who were unrelated individuals without type 2 diabetes. Diagnosis criteria of type 2 diabetes mellitus were based on American Diabetes Association, with fasting glucose $\geq$ 126 mg/dL, 2 h plasma glucose $\geq$ 200 mg/dL during glucose tolerance test. Before recruiting in the study, each individual was informed

about the viewpoint of this study and after getting their full consent (written); they were included in the study. A structured questionnaire was used to elicit detailed information. All individuals were interviewed and information regarding their age, gender, type 2 diabetes medication use, presence of any complication, history of other disorders, and smoking habit was collected questionnaire to match cases and controls for both ethnicities by trained research staff. The study was approved by the Ethical Review Board of the Faculty of Biological Sciences, University of Dhaka, Bangladesh. Five mL of venous blood was collected into EDTA containing vacutainer tubes from each individual by a certified phlebotomist using standard laboratory techniques. Genomic DNA was extracted by organic extraction method from the cellular fraction of blood as discussed in our previous manuscript (Afruza et al. 2014). The quality and quantity of extracted DNA was measured using NanoDrop (NanoDrop1000, US). The isolated DNA was stored at − 20 °C until genotyping.

Power of the Genetic Association Study

The power of the proposed study is calculated using the Genetic Association Study (GAS) Power Calculator which is based on the CaTS power calculator for two-stage association studies (Skol et al. 2006). With the 217 T2DM patients and 237 healthy controls, 12% disease prevalence, 10% minor allele frequency in cases, and a genotype relative risk of 1.8 with 5% significance level, gives a power of the test to be 93%.

Molecular Genetic Study

The microRNA-binding site polymorphism, rs1366600, was genotyped by using the TaqMan SNP genotyping assay (Reference ID: C_8356110_20) based on real time PCR based allelic discrimination method. The assay contained the two primers and allele-specific fluorescent probes labeled by FAM and VIC, respectively. Genomic DNA was diluted to 10 ng/µL. The reaction mixture of PCR was 10 µL comprising of 2.5 µL TaqMan genotyping master mix (2×), 0.13 µL of TaqMan genotyping assay mix, 2.625 µL nuclease-free water and 4.7 µL of the diluted DNA sample. A negative control was used containing diH_2O in place of genomic DNA. The PCR condition included an initial hot start step at 95 °C for 10 min and the 35 cycles of denaturation at 95 °C for 15 s and annealing at 60 °C for 60 s. The signal intensity was measured in the Applied Biosystems 7500 real time PCR instrument. Genotyping or allele calling was carried out with Applied Biosystems 7500 software version 2.0.5.

Statistical Analyses

The genotype and allelic frequency of single nucleotide polymorphism rs1366600 was calculated using the SNP_tools add-Ins program of Excel (Chen et al. 2009). The allele frequency was calculated in the following way: one individual with homozygous TT genotype has 2 T-alleles, TC genotype contains one T-allele and

one C-allele, and CC genotype contains 2 C-alleles. Thus, the frequency of the T-allele is: (no of T-alleles/Total no. of allele). The frequency of the C-allele is: (no of C-alleles/total no. of allele). The calculation of odds ratio, P-values, and the Chi square tests were also performed using the SNP_tools. The genetic association was calculated using additive, dominant, and recessive model of inheritance.

Results

Retrieval of SNPs Data

The Ensembl genome browser 88 (March, 2017) was searched for retrieving the SNPs in the human INSR gene and from SNP's data missense variant, 5′ UTR variant, and 3′ UTR variant were selected. The total number of SNPs for INSR gene was 12652. The global MAF of these SNPs ranged from 0.1 to 49.9%. Further, we wanted to check the relative distribution of these SNPs across different categories. Since, the nsSNPs are functionally significant SNPs (altering amino acid that they code for) and the UTR SNPs by virtue of their residence in the regulatory region (miRNA or transcription factor binding sites) are also functionally relevant. Among all the INSR SNPs, 429 (3.39%) were nsSNPs, 6 (0.04%) were in 5′ UTR region, and 54 (0.42%) were in 3′ UTR region (Fig. 1). Our ultimate target was to further analyze the miRNA binding site polymorphisms located at the 3′-UTR region of a gene.

Functional Significance of the 3′ UTR SNPs

The 3′ UTR SNPs of the *INSR* gene were analyzed for their possible functional impact using the RegulomeDB database. The results from our analyses showed that the SNPs could be put into different groups with potential influence of transcription

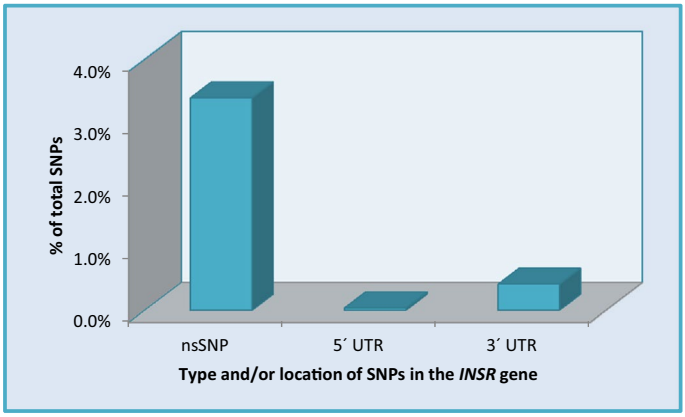

Fig. 1 Distribution of nonsynonymous (ns), 5′ UTR and 3′ UTR SNPs for INSR gene (based on the Ensembl genome browser 88)

factor (TF) binding in virtue of their location within the TF binding motif and/or their residence in the DNase hypersensitivity sites as indicated by the DNase peak. The results of RegulomeDB for these SNPs are shown in Table 1.

Additionally, since the 3′ UTR region is a potential site for miRNA binding, we used PolymiRTS database to predict whether the SNPs in the 3′ UTR region of the *INSR* gene change the binding motif for miRNAs. Our analysis showed that the rs1366600 SNP is most significant SNPs; whose both alleles are associated with predicted disruption (functional class 'D') and creation (functional class 'C') of binding motifs for a number of miRNAs (Table 1).

Rationale of Choosing rs1366600 for Case–Control Study

The SNP rs1366600 located in the insulin receptor gene (*INSR*) was found as an important candidate SNP for association with T2DM in different population (Zhao et al. 2013; Wang et al. 2017). rs1366600 is located in a microRNA-binding site in the 3′-UTR region of the *INSR* gene. However, there is no such study in Bangladeshi population. For this reason, we selected rs1366600 as a candidate SNP and validated it in Bangladeshi population for potential association with T2DM.

Genetic Association Study of rs1366600 T > C with the Risk of T2DM in Bangladeshi Individuals

To further validate the findings of our in silico analysis of rs1366600 as a potential SNP altering miRNA binding and thus increasing the susceptibility of disease, we carried out a case control study. The study was carried out to evaluate the association of genetic polymorphism (rs1366600) residing in the miRNA binding site of the *INSR* gene with the risk of T2DM in Bangladeshi individuals.

Demographic, Anthropometric and Clinical Data of T2DM Patients and Healthy Controls

In this study a total of 454 Bangladeshi individuals were included and among them 226 were male and 228 were female. The number of T2DM patients was 217 (male = 95; female = 122) and the number of healthy controls was 237 (male = 131; female = 106). The data for the average age, weight, height, and BMI of T2DM patients and healthy controls were collected. Blood pressure, fasting plasma glucose (FPG), and hemoglobin A1c (HbA1C) levels of the T2DM patients and healthy controls were also measured. The demographic, anthropometric, and clinical data of T2DM patients and healthy controls included in this study are shown in Table 2.

Genotype Assignment for the rs1366600 T > C SNP from the Allelic Discrimination Assay

The analysis of the results obtained from the allelic discrimination plots showed that 145 T2DM patients and 190 healthy controls carry the wildtype TT genotype.

Table 1 Linking 3′ UTR SNPs of *INSR* gene with effect on miRNA binding as predicted by PolymiRTS and their functional significance analyzed by RegulomeDB

dbSNP ID	Ancestral allele	Allele	miR ID	Functional Class[a]	RegulomeDB score[b]	Supporting data
rs1366600	**T**	**T**	hsa-miR-106a-5p, hsa-miR-106b-5p, hsa-miR-17-5p, hsa-miR-20a-5p, hsa-miR-20b-5p, hsa-miR-302a-3p, hsa-miR-302b-3p, hsa-miR-302c-3p, hsa-miR-302d-3p, hsa-miR-302e, hsa-miR-372-3p, hsa-miR-373-3p, hsa-miR-512-3p, hsa-miR-519d-3p, hsa-miR-520a-3p, hsa-miR-520b, hsa-miR-520c-3p, hsa-miR-520d-3p, hsa-miR-520e, hsa-miR-526b-3p, hsa-miR-548az-5p, hsa-miR-548t-5p, hsa-miR-93-5p	D	**3b**	**TF binding + matched TF motif**
		C	hsa-miR-18a-5p, hsa-miR-18b-5p, hsa-miR-4735-3p	C		
rs192789895	C	T	hsa-miR-448	C	4	TF binding + DNase peak
rs116953519	C	T	hsa-miR-190a-5p, hsa-miR-190b, hsa-miR-3689a-5p, hsa-miR-3689b-5p, hsa-miR-3689e, hsa-miR-3689f,	C	4	TF binding + DNase peak
rs115358150	A	A	hsa-miR-3174, hsa-miR-4524a-3p, hsa-miR-921	D	4	TF binding + DNase peak
		G	hsa-miR-103b	C		
rs138486944	G	G	hsa-miR-21-5p, hsa-miR-590-5p, hsa-miR-634	D	4	TF binding + DNase peak
rs192753442	A	A	hsa-miR-1271-3p, hsa-miR-550a-3-5p, hsa-miR-550a-5p, hsa-miR-550b-2-5p, hsa-miR-6500-3p	D	4	TF binding + DNase peak
		G	hsa-miR-668-5p, hsa-miR-6720-3p	C		
rs146905190	G	G	hsa-miR-6734-3p	D	4	TF binding + DNase peak
		A	hsa-miR-1343-3p, hsa-miR-221-5p, hsa-miR-3667-3p, hsa-miR-6783-3p, hsa-miR-8073, hsa-miR-874-5p	C		

[a]**Functional class: C**: The derived allele creates a new miRNA site; D: The derived allele disrupts a conserved miRNA site

[b]**RegulomeDB score: 3b**: less likely to affect binding; 4: minimal binding evidence

Table 2 Demographic, anthropometric and clinical data of T2DM patients and healthy controls

Parameters	T2DM patients ($n=217$)	Healthy controls ($n=237$)	P value (t test)
Gender			
Male (n)	95	131	Not calculated
Female (n)	122	106	
Age (years), mean(±SD)	50.19 (±11.44)	42.18 (±16.33)	<0.0001
Weight (kg), mean(±SD)	62.49 (±6.48)	64.11 (±10.25)	0.05
Height (cm), mean(±SD)	163.03 (±3.11)	164.68 (±7.17)	0.0019
BMI (kg/m^2), mean(±SD)	23.52 (±2.46)	23.69 (±3.85)	0.58
Fasting plasma glucose, FPG (mmol/L)	9.13 (±1.10)	4.31 (±1.20)	<0.0001
HbA1C (%)	10.02 (±2.26)	4.50 (±1.07)	<0.0001

However, the minor allele was present at a low frequency in its homozygous form in the study population (the number of T2DM patients with CC genotypes are 8 and for the healthy controls it is 4). The number of T2DM patients with the TC heterozygous genotype was 64 and 43 of the healthy controls carried TC genotype.

Genotypic and Allelic Frequency Distribution for the rs1366600 (T > C) Polymorphism in T2DM Patients and Healthy Controls

The frequency of the TT genotype of the rs1366600 (T > C) SNP is higher than TC and CC genotypes in both the T2DM patients and healthy controls from Bangladesh. The genotypic and allelic frequency distribution for the rs1366600 (T > C) polymorphism in Bangladeshi T2DM patients and healthy controls are shown in Table 3.

Table 3 Genotypic and allelic frequency distribution for the rs1366600 (T > C) polymorphism in T2DM patients and healthy controls

Study subject	Frequency and samples number	Genotypic frequency for rs1366600 (T > C)			Allelic frequency	
		Wild type homozygous	Rare homozygous	Heterozygous	Wild type allele	Minor allele
		TT	CC	TC	T	C
T2DM patients	Frequency	0.67	0.04	0.29	0.82	0.18
	n	145	8	64	354	80
Healthy controls	Frequency	0.80	0.02	0.18	0.89	0.11
	n	190	4	43	423	51

Genetic Association Analysis of the rs1366600 (T > C) Polymorphism with T2DM Patients and Healthy Controls in Bangladeshi Population

Logistic regression analysis showed that the minor allele (C-allele) of the rs1366600 (T > C) polymorphism is associated with increased risk of T2DM in Bangladeshi individuals (Odds ratio (OR) 1.87; 95% CI 1.28–2.74; $P = 0.0010$). When we dissected our analysis to include the dominant model (CC + TC genotypes against the wild type TT genotype), we found that the CC and TC genotypes; in combination were associated with increased risk of T2DM in Bangladeshi population (OR 2.01; 95% CI 1.31–3.07; $P = 0.0012$). However, in recessive model (CC vs TT + TC); the effect was not statistically significant (OR 2.23; 95% CI 0.66–7.51; $P = 0.1848$). The results of logistic regression analysis are shown in Table 4.

Further, we stratified our analysis based on the gender of the individuals. Our analysis showed that the association of the polymorphism rs1366600 (T > C) with the risk of T2DM in Bangladeshi male individuals remain similarly significant across different models (Table 5). The female population did not show statistically significant association (Table 5).

Discussion

To the best of our knowledge, this is the first study of blending in silico analysis with case–control studies for analyzing genetic association of T2DM patients in Bangladesh. The principal objectives of this study were to first employ in silico approaches to characterize SNPs harbored in the diabetes associated genes that are functionally deleterious and further investigate the association of one of the potential SNPs with type 2 diabetes mellitus in Bangladeshi population using laboratory based case–control approaches.

Table 4 Genetic association analysis of the rs1366600 (T > C) polymorphism with T2DM patients and healthy controls in Bangladeshi population

Model	Genotypes/alleles	T2DM patients	Healthy controls	OR (95% CI)	*P* value
Additive	T	354	423	Reference	**0.0010**
	C	80	51	1.87 (1.28–2.74)	
	TT	145	190	Reference	**0.0029**
	TC	64	43	1.95 (1.25–3.04)	
	TT	145	190	Reference	0.1089
	CC	8	4	2.62 (0.77–8.87)	
Dominant	TT	145	190	Reference	**0.0012**
	CC + TC	72	47	2.01 (1.31–3.07)	
Recessive	TT + TC	209	233	Reference	0.1848
	CC	8	4	2.23 (0.66–7.51)	

Bold values indicate statistically significant association

Table 5 Impact of gender for the association of the rs1366600 (T > C) polymorphism with T2DM patients and healthy controls in Bangladeshi population

Gender	Model	Genotypes/alleles	T2DM patients	Healthy controls	OR (95% CI)	*P* value
Male Individuals	Additive	T	151	234	Reference	**0.0037**
		C	39	28	2.16 (1.27–3.66)	
		TT	61	106	Reference	**0.0098**
		TC	29	22	2.29 (1.21–4.33)	
		TT	61	106	Reference	0.1387
		CC	5	3	2.90 (0.67–12.54)	
	Dominant	TT	61	106	Reference	**0.0048**
		CC + TC	34	25	2.36 (1.29–4.33)	
	Recessive	TT + TC	90	128	Reference	0.2325
		CC	5	3	2.37 (0.55–10.17)	
Female undividuals	Additive	T	203	189	Reference	
		C	41	23	1.66 (0.96–2.87)	0.0679
		TT	84	84	Reference	
		TC	35	21	1.67 (0.90–3.10)	0.1045
		TT	84	84	Reference	
		CC	3	1	3.00 (0.31–29.43)	0.3230
	Dominant	TT	84	84	Reference	
		CC + TC	38	22	1.73 (0.94–3.17)	0.0755
	Recessive	TT + TC	119	105	Reference	
		CC	3	1	2.65 (0.27–25.84)	0.3846

Bold values indicate statistically significant association

After computationally validating rs1366600 as a potential functional SNP in the *INSR* gene, we studied association of this polymorphism with Type 2 diabetes mellitus (T2DM) in Bangladesh. In the present case–control study, considering small deviation from different genetic models of insulin receptor gene polymorphism, we investigated the associations of rs1366600 with T2DM risk based on additive, dominant, and recessive models. We found a significant result under the additive model, which could be a reason of affected genotypes and insufficient statistical power in some models. The lower number of CC genotypes in both T2DM cases and controls in the cohort could affect the overall power of the additive model to detect true association. However, the risk association is seen in the dominant model, when the CT

and CC genotypes are pooled together for association analysis. The recessive model is clearly lacking enough number of cases and controls to describe true association.

Among the predicted high impact functional SNPs in the *INSR* gene, rs1366600 was found to be significantly associated with T2DM and gestational diabetes in Chinese population (Zhao et al. 2013; Wang et al. 2017). rs1366600 is a micro RNA binding site polymorphism. miRNA binding site SNPs are considered goldmine for the genetic epidemiological studies and assumed to contribute to the susceptibility of multiple human diseases (Chen et al. 2008). Studies have also demonstrated that a single nucleotide mismatch between the seed region and its target site can abolish repression (Brennecke et al. 2005; Brodersen and Voinnet 2009), thus establishing the requirement of stringent recognition between the two interacting sequences. According to previous studies, the 3′ UTR SNP located in the microRNA-binding site of *HNF1B, WFS1 and various other genes* were found to be associated with T2DM susceptibility in different population (Gong et al. 2014; Elek et al. 2015; Goda et al. 2015b). Moszyńska A et al. reviewed the impacts of 3′ UTR SNPs in human diseases ranging from cancer to diabetes and presented that SNPs affecting microRNA-binding sites in the 3′ UTR regions can lead to disease pathogenesis via altering mRNA stability (Moszyńska et al. 2017). However, no studies have reported any association of microRNA-binding site polymorphism in Bangladeshi T2DM individuals.

The global minor allele frequency (MAF) reported in the dbSNP database is 8.35% (Sherry et al. 2001). In the Ensembl database (Zerbino et al. 2018), the MAF for this SNP is 8%. However, the frequency of the minor allele 'C' is very much different in different population. According to the 1000 Genomes Project Phase 3 allele frequencies, the highest MAF for rs1366600 is 22% in the African population and the lowest is 0.1% in the European population. The Asian population has a MAF of 6.3% (The Genomes Project 2015). In a similar study on the association rs1366600 with the risk of T2DM in Chinese Han population, it was shown that the MAF of this SNP is 15% in T2DM cases ($n = 1008$) and 12.5% in controls ($n = 1052$) (Zhao et al. 2013). In the current study, we found that rs1366600 has a MAF of 18% in cases and 11% in controls.

In concordance with the studies carried out in different population, we could find an increased risk for the susceptibility of T2DM with the minor allele and genotypes of the rs1366600 SNP in Bangladeshi population. Our results showed that the minor allele of the microRNA-binding site polymorphism rs13666000 is associated with almost two-fold increased risk of T2DM, compared to the individual not carrying the minor allele 'C'. Dissecting our analysis into different models of inheritance pattern showed that the dominant model (CC + TC genotype against the TT genotype), confers two fold increased susceptibility to T2DM in Bangladeshi population. The value and novelty of this study are identification of candidate SNP by in silico method and validation of it by genetic association study.

In order to gain more insight into the functionality of rs1366600, we analyzed for the eQTL values in the GTEx project portal (Keen and Moore 2015). However, we could not find any eQTL data for all tissue GTEx. But for single tissue analysis, separately, we found some data on different tissues which are insignificant or showed that the allelic variations in rs1366600 are not affecting any tissue expression (data

not shown). INSR is expressed in more than 25 different tissues and the highest expression is in Spleen and Kidney (Fagerberg et al. 2014). However, Kidney has not any GTEx data for rs1366600.

In conclusion, our results suggest the rs1366600 located in the microRNA-binding site of the *INSR* gene shows significant association with T2DM risk in Bangladeshi population. Further, functional studies are necessary to investigate whether the miRNA regulates the expression of target gene by binding to predicted target sequences.

Acknowledgement We would like to acknowledge the generosity of the individuals who participated in this study. We would also like to thank Professor Dr Haseena Khan, Department of Biochemistry and Molecular Biology, University of Dhaka for providing us the facilities for performing RT-PCR based allelic discrimination assay.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with Ethical Standards

Conflict of Interest The authors declare no conflict of interest.

References

Afruza R, Islam LN, Banerjee S, Hassan MM, Suzuki F, Nabi AN (2014) Renin gene polymorphisms in bangladeshi hypertensive population. J Genomics 2:45–53

Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, García Girón C, Hourlier T, Howe K, Kähäri A, Kokocinski F, Martin FJ, Murphy DN, Nag R, Ruffier M, Schuster M, Tang YA, Vogel J-H, White S, Zadissa A, Flicek P, Searle SMJ (2016) The Ensembl gene annotation system. Database, baw093

Akter S, Rahman MM, Abe SK, Sultana P (2014) Prevalence of diabetes and prediabetes and their risk factors among Bangladeshi adults: a nationwide survey. Bull World Health Organ 92(204–13):213a

Bhattacharya A, Ziebarth JD, Cui Y (2014) PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. Nucleic Acids Res 42:D86–D91

Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M (2012) Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22:1790–1797

Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3:e85

Brodersen P, Voinnet O (2009) Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol 10:141–148

Chen K, Song F, Calin GA, Wei Q, Hao X, Zhang W (2008) Polymorphisms in microRNA targets: a gold mine for molecular epidemiology. Carcinogenesis 29:1306–1311

Chen B, Wilkening S, Drechsel M, Hemminki K (2009) SNP_tools: a compact tool package for analysis and conversion of genotype data for MS-Excel. BMC Res Notes 2:214

Chowdhury MA, Uddin MJ, Khan HM, Haque MR (2015) Type 2 diabetes and its correlates among adults in Bangladesh: a population based study. BMC Public Health 15:1070

De Silva NM, Frayling TM (2010) Novel biological insights emerging from genetic studies of type 2 diabetes and related metabolic traits. Curr Opin Lipidol 21:44–50

Elek Z, Nemeth N, Nagy G, Nemeth H, Somogyi A, Hosszufalusi N, Sasvari-Szekely M, Ronai Z (2015) Micro-RNA binding site polymorphisms in the WFS1 gene are risk factors of diabetes mellitus. PLoS ONE 10:e0139519

Fagerberg L, Hallstrom BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjostedt E, Lundberg E, Szigyarto CA, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson A, von Feilitzen K, Forsberg M, Zwahlen M, Olsson I, Navani S, Huss M, Nielsen J, Ponten F, Uhlen M (2014) Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics 13:397–406

Goda N, Murase H, Kasezawa N, Goda T, Yamakawa-Kobayashi K (2015a) Polymorphism in microRNA-binding site in HNF1B influences the susceptibility of type 2 diabetes mellitus: a population based case-control study. BMC Med Genet 16:75

Goda N, Murase H, Kasezawa N, Goda T, Yamakawa-Kobayashi K (2015b) Polymorphism in microRNA-binding site in HNF1B influences the susceptibility of type 2 diabetes mellitus: a population based case–control study. BMC Med Genet 16:75

Gong W, Xiao D, Ming G, Yin J, Zhou H, Liu Z (2014) Type 2 diabetes mellitus-related genetic polymorphisms in microRNAs and microRNA target sites. J Diabetes 6:279–289

Grarup N, Sandholt CH, Hansen T, Pedersen O (2014) Genetic susceptibility to type 2 diabetes and obesity: from genome-wide association studies to rare variants and beyond. Diabetologia 57:1528–1541

Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE (2014) Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes Res Clin Pract 103:137–149

He Y, Ding Y, Liang B, Lin J, Kim TK, Yu H, Hang H, Wang K (2017) A systematic study of dysregulated microRNA in type 2 diabetes mellitus. Int J Mol Sci 18:456

Kahn SE, Cooper ME, Del Prato S (2014) Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. Lancet 383:1068–1083

Keen JC, Moore HM (2015) The genotype-tissue expression (GTEx) project: linking clinical data with molecular analysis to advance personalized medicine. J Pers Med 5:22–29

Lyssenko V, Groop L (2009) Genome-wide association study for type 2 diabetes: clinical applications. Curr Opin Lipidol 20:87–91

Moszyńska A, Gebert M, Collawn JF, Bartoszewski R (2017) SNPs in microRNA target sites and their potential role in human disease. Open Biology 7:170019

Pradeepa R, Mohan V (2017) Prevalence of type 2 diabetes and its complications in India and economic costs to the nation. Eur J Clin Nutr 71:816–824

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–311

Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 38:209

The Genomes Project, C. (2015) A global reference for human genetic variation. Nature 526:68

Wang X, Li W, Ma L, Ping F, Liu J, Wu X, Mao J, Wang X, Nie M (2017) Investigation of miRNA-binding site variants and risk of gestational diabetes mellitus in Chinese pregnant women. Acta Diabetol 54:309–316

Whiting DR, Guariguata L, Weil C, Shaw J (2011) IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Res Clin Pract 94:311–321

Wu B, Miller D (2017) Involvement of microRNAs in diabetes and its complications. Methods Mol Biol 1617:225–239

Wu Y, Ding Y, Tanaka Y, Zhang W (2014) Risk factors contributing to type 2 diabetes and recent advances in the treatment and prevention. Int J Med Sci 11:1185–1200

Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG, Gil L, Gordon L, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, To JK, Laird MR, Lavidas I, Liu Z, Loveland JE, Maurel T, McLaren W, Moore B, Mudge J, Murphy DN, Newman V, Nuhn M, Ogeh D, Ong CK, Parker A, Patricio M, Riat HS, Schuilenburg H, Sheppard D, Sparrow H, Taylor K, Thormann A, Vullo A, Walts B, Zadissa A, Frankish A, Hunt SE, Kostadima M, Langridge N, Martin FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Aken BL, Cunningham F, Yates A, Flicek P (2018) Ensembl 2018. Nucleic Acids Res 46:D754–D761

Zhao X, Ye Q, Xu K, Cheng J, Gao Y, Li Q, Du J, Shi H, Zhou L (2013) Single-nucleotide polymorphisms inside microRNA target sites influence the susceptibility to type 2 diabetes. J Hum Genet 58:135–141

Zhou X, Yang PC (2012) MicroRNA: a small molecule with a big biological impact. Microrna 1:1

Affiliations

Mahrima Parvin[1] · Farhana Jahan[1] · Pankaj Kumar Sarkar[2] · Zakir Hossain Howlader[3] · A. H. M. Nurun Nabi[1] · Md. Ismail Hosen[1,3]

[1] Laboratory of Population Genetics, Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka 1000, Bangladesh

[2] Dinajpur Diabetes O Swasthoseba Hospital, Dinajpur, Bangladesh

[3] Nutritional Biochemistry Laboratory, Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka, Bangladesh