

# The AGT Epistasis Pattern Proposed a Novel Role for ZBED9 In Regulating Blood Pressure: Tehran Cardiometabolic Genetic Study (TCGS)

**Mahdi Akbarzadeh**

Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

**Parisa Riahi**

Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

**Goodarz Kolifarhood**

Department of Pain Genomics, McGill University, QC

**Hossein Lanjanian**

Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

**Nadia Alipour**

Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University

**Leila Najd Hassan Bonab**

Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

**Mohammad Reza Moghadas**

Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

**Siamak Sabour**

Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran

**Fereidoun Azizi**

Endocrine research center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

**Maryam S Daneshpour (✉ [daneshpour@sbmu.ac.ir](mailto:daneshpour@sbmu.ac.ir))**

Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran

---

**Research Article**

**Keywords:** Hypertension, SKAT, Mb-MDR, ZBED9, Epistasis, AGT, TNXB

**Posted Date:** January 11th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1216294/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

**Background:** Hypertension is typically considered as the leading risk factor for cardiovascular disease. Epistasis studies may add another layer of complexity to our understanding of the genetic basis of hypertension.

**Methods:** A nested case-control design was used on 4214 unrelated Tehran Cardiometabolic Genetic Study (TCGS) adults to evaluate 65 SNPs of previously associated genes, including ZBED9, AGT, and TNXB. The integrated effect of each gene was determined using the Sequence-based Kernel Association Test (SKAT). We used model-based multifactor dimension reduction (Mb-MDR) and entropy-based gene-gene interaction (IGENT) methods to determine interaction and epistasis patterns.

**Results:** The integrated effect of each gene has a statistically significant association with blood pressure traits ( $P$ -value  $< 0.05$ ). Single-locus analysis identified two missense variants in ZBED9 (rs450630) and AGT (rs4762) that are associated with hypertension. In the ZBED9 gene, significant local interactions were discovered. The G allele in rs450630 showed an antagonistic effect on hypertension, but interestingly, IGENT analysis revealed significant epistasis effects for different combinations of ZBED9, AGT, and TNXB loci.

**Conclusion:** We discovered a novel interaction effect between a significant variant in an essential gene for hypertension (AGT) and a missense variant in ZBED9, which has shifted our focus to ZBED9's role in blood pressure regulation.

## 1. Introduction

Hypertension (HTN), as a silent killer, is one of the leading causes of death due to cardiovascular diseases(1). Globally, a number of 1.38 billion people with high blood pressure (systolic blood pressure (SBP)  $\geq 140$  mmHg and/or diastolic blood pressure (DBP)  $\geq 90$  mmHg) had been estimated in adults age over 18 years worldwide in 2010 (Mills et al. 2016; Mills, Stefanescu, and He 2020). However, the incidence and prevalence of HTN have been distributed disproportionately with a high burden of disease in low- to middle-income countries (4).

Epidemiological findings had stressed the fact that high blood pressure is primarily due to ageing, physical inactivity and unhealthy diets including, high sodium and low potassium intake, in the presence of a genetic predisposition (Mills et al. 2016; Mills, Stefanescu, and He 2020). As a complex disease, at least 30% of SBP and DBP variances are attributed to genetic factors in family-based heritability analysis (5). Further, genome-wide association studies (GWAS) have identified a large number of loci that are associated with blood pressure traits in diverse populations, of which the majority of genetic variants have been reported in individuals with European Ancestry (6).

Iran, as a middle-income country, is forced to halt the increasing rate of hypertension. According to national reports, almost half of Iranian adults live with high blood pressure (7,8). Similar to GWAS findings in the European population, *AGT* and *TNXB* with a consistent effect on SBP, DBP and HTN are contributed to blood pressure regulation among Iranian families(9). Further, a recent genome-wide association study in the Iranian population discovered a new locus (*ZBED9*) for hypertension that was confirmed in a linkage analysis on different sets of blood pressure traits and replicated in summary statistics of UK Biobank on DBP and HTN (10). Accordingly, the current study aimed to approve epistasis effects of *AGT*, *ZBED9*, and *TNXB* on blood pressure traits using different statistical approaches, including non-parametric machine-learning algorithm, Model-based Multifactor Dimensionality Reduction (Mb-MDR), and entropy-based MDR interaction network.

## 2. Methods And Materials

### 2.1. Study participants and study design

Tehran Cardiometabolic Genetic Study (TCGS) is a family-based genetic database according to a larger study, Tehran Lipid and Glucose Study (TLGS)(11–13). TLGS is a population-based cohort started in 1999, and study subjects have been followed up through six phases (every three years) on non-communicable disorders' (NCDs) and related risk factors (14). The current study extracted 4,214 unrelated participants aged 18 or above, based on the well-designed protocol in previous GWAS in TCGS (9,10). HTN incident cases and a random sample of healthy individuals with two or more follow-up records were included in the study using a nested case-control sampling design.

### 2.2. Ethics and consent

This research was approved by the Medical Ethics Committee of Shahid Beheshti University of Medical Sciences and the National Committee for Ethics in Biomedical Research in May 2021 with code IR.SBMU.ENDOCRINE.REC.1400.024. All participants gave written informed consent to participate. The research has been performed in accordance with the Declaration of Helsinki.

## 2.3. Measurements

Principles for clinical investigations could be found in the original TCGS and TLGS papers (19,22). Having SBP  $\geq 140$  mmHg or DBP  $\geq 90$  mmHg or taking antihypertensive treatment was hypertension definition. Risk factors of high blood pressure traits (SBP, DBP, and HTN) were chosen based on previous GWAS literature(6,9,10,15). The following predictors were included in the model as covariates: sex, age, and body mass index (BMI). BMI was measured as the ratio of body weight to squared height ( $\text{kg}/\text{m}^2$ ). Using the Amelia package in R, height and weight were imputed multiple times using the Expectation-Maximization method with Bootstrapping (EMB) approach, for which the missing rate of covariate's values ranged from 0.2–7.7% over five phases of TLGS. Traits and their risk factors were tested for their distribution among males and females, using the T-test and Chi-square test.

## 2.4. Genotyping and quality control

Blood samples of TCGS participants were genotyped using humanOmniExpress-24-v1 bead chips on 652919 single nucleotide polymorphisms with an average mean distance of 4 kilobases for each individual at deCODE genetic company (16). Genomic quality control (QC) was implemented in comprehensive steps(10,17).

## 2.5. Testing the integrated effect of gene

Sequence Kernel Association Test (SKAT) was used to test the overall effect of *AGT*, *ZBED9*, and *TNXB*. Optimal SNP-set kernel-based association test (SKAT-O) is used to test the association between a set of variants in a region in the genome, either rare or common or both cumulative effects, and a trait (binary or quantitative) by a linear or logistic regression framework(18). SKAT-O obtains an aggregated score of individual variants and efficiently obtains a corresponding SNP-set *P-value*. Moreover, the covariates are adjusted in SKAT-O. In the case of the linear kernel, test statistic can be expressed as a weighted sum of the single-marker score test statistic. SKAT-O is usually considered as a variance component test that assumes the genotyping effects are random. *P-value* is provided for each set of variants of *ZBED9*, *AGT* and *TNXB*, separately. All analyses are done using the *SKAT* R package, version 2.0.1 (19). Readers can find the workflow of gene-gene interactions between *AGT*, *ZBED9*, and *TNXB* involved in high blood pressure traits (SBP, DBP, and hypertension) in Supplementary Figure. 1

## 2.6. Making variant collection

After specifying genes, 65 variants, including three previously reported tag SNPs (*ZBED9*: rs450630, *TNXB*: rs2021783, *AGT*: rs2493134), were extracted in the TCGS database. Next, variants in the selected regions that were in low to moderate LD ( $r^2 \leq 0.4$ ) with the tag SNPs were extracted ( $N=13$ ). The relief algorithm, as a machine learning approach for classification and feature selection, was applied to extract the most significant variants to reduce the computational expenses (20). The final set consisted of 11 SNPs, including four variants located in *TNXB* and *AGT* and three in *ZBED9*. Supplementary Table. 1 shows the information of genetic variants in *ZBED9*, *AGT*, and *TNXB* and their LDs. Figure 1 depicts the three genes. As illustrated in the figure, the *ZBED9* gene contains two polymorphisms in exons 3 and 4, and one polymorphism in intron region 2. One polymorphism is located in exon 21 and the other three are located in the introns of the *TNXB* gene. Additionally, the *AGT* gene contains a variant in exon 2 and others in the intronic regions, and the interaction of these polymorphisms results in a significant association with hypertension.

## 2.7. Single Locus Analysis

The generalized Linear Model (GLM) was employed in four different genetic models to determine the optimal genetic model for every single variant and extract the effects of causal variants. These models included additive, dominant, recessive, and co-dominant models. To investigate the effect size of rare variants ( $\text{MAF} < 0.05$ ), Firth's Bias-reduced logistic regression was used by the R *logistf* package (21). Coding labels for the additive model were defined based on the minor allele count, as BB=0, AB=1, and AA=2. Dominant, recessive and co-dominant models were defined as AA + AB versus BB, AA versus AB+BB and AB versus AA+BB genotypes, respectively.

## 2.8. Epistasis and Local Interactions

Model-based Multifactor Dimensionality reduction (MB-MDR) has been proposed by Calle et al. as a semiparametric machine learning dimension-reduction algorithm. MB-MDR incorporates three main steps, (i) constructing two-way contingency tables for the

combination of SNPs (p-locus model) and creating a new categorical variable X. (ii) computing the wald statistic for the testing association of each X variable and the trait. (iii) conducting a permutation test to prevent inflating alpha error rate (Calle et al. 2008). In the first step, two-way contingency tables are constructed for each SNP combination (p-locus model). Each cell in the contingency table is considered as a “Low Risk”, “High-Risk”, or “No-Effect” category, based on the student’s t-test statistic, to test if the mean difference of a cell with the rest of the cells is either significant or non-significant. If the difference between cell means is significant at  $P$ -value=0.10, the cell will be labelled as either “Low Risk” or “High-Risk”, based on the t statistic’s sign; otherwise, it will be considered as “No-Effect”. If the sign of the t statistic is positive, then the corresponding cell will be labelled as “High-Risk”; otherwise, it is “Low Risk”. Thus, a new “categorical variable X= {“No effect”, “Low effect”, “High effect”}” will be created in the first step of Mb-MDR. The Wald test statistic is conducted between the categorical variable X and the trait in the second step. We considered the maximum of “wald-low effect” and “wald-high effect”, which are statistics for the mean differences of the “Low effect” and “High effect” category with two other categories, respectively. To avoid increasing type I error, a permutation test and bootstrapping as cross-validation tests are being assessed for evaluating the significance of the maximum of “wald-low effect” and “wald-high effect” for each p-locus model in the third step (Calle et al. 2008).

The current study employed this method to detect antagonistic and synergistic interactions within (local) and between loci of *ZBED9* with *AGT* and *TNXB*. All interaction effects were calculated by the “*mbmdr*” package in R version 4.0.3 software, and  $P$ -values were adjusted using the Bonferroni correction method. Multiple logistic regression was employed to capture the corresponding significant genotypes for each interaction model, both synergistic and antagonistic interactions. To obtain an entropy-based interaction graph and visualize pairwise interactions between variants, *mdr* 3.0.2 software was used. To do so, independent effects of variants and their pairwise interaction effects were assessed and visualized using IGENT, entropy-based gene-gene interaction method for genome-wide interaction analysis. Fully description of the IGENT is described elsewhere (23).

## 2.9 Investigation of the missense variations affect the proteins

We applied the bioinformatics analysis to investigate the effects of the missense variation on the structure of the proteins coded by our interesting genes. The ZBED9 is an intrinsically disordered protein (IDP) [1, 2]. In the free and native conditions, IDPs have no single well-defined tertiary structure. Therefore, the bioinformatics tools were applied to identify disordered regions and disordered binding regions of this protein. IUPred2A [3, 4] is a valid and user-friendly webserver to Predict our desired information. The sequence of ZBED9 was obtained from NCBI with the protein Accession Q6R2W3.1. The 6i3f PDB file [5] of the RCSB database was used as the 3D structure of the AGT protein. The visualization of PDB files has been done by Chimera [6].

## 3. Results

### 3.1. Basic description of genotyped samples

Supplementary Table. 2 presents the distribution of demographic and clinical characteristics of 4214 independent participants by sex with a similar distribution of age and significant differences in BMI, SBP, DBP, and HTN. Two variants (rs3130287(C>T) and rs11969759(C>T) were excluded by reliefF algorithm, and a set of intron, UTR and missense variants on *ZBED9*, *AGT* and *TNXB* were considered in the final analysis. The majority of variants were in low to moderate LD ( $r^2$ : 0-34%). Except for rs2021783 on *AGT* and rs449074 on *ZBED9*, remained variants were common with MAF>5%.

### 3.2. SKAT-O

The genetic variants on corresponding genes were treated as an individual gene set, consisted of rare and common variants, and were considered in the SKAT-O to check for aggregated association with blood pressure traits. All of the gene sets were associated with hypertension (*ZBED9*:  $P$ -value= 0.0268, *AGT*:  $P$ -value=0.0402, *TNXB*:  $P$ -value= 0.0035), SBP (*ZBED9*:  $P$ -value= 0.0186, *AGT*:  $P$ -value=0.0038, *TNXB*:  $P$ -value= 0.0021), and DBP (*ZBED9*:  $P$ -value= 0.006, *AGT*:  $P$ -value=0.010, *TNXB*:  $P$ -value= 0.028).

### 3.3. Single locus analysis

GLM results for single locus analysis are described in detail in Supplementary Table 3. The association of rs450630 on *ZBED9* and HTN was significant under the recessive (OR=1.35,  $P$ -value=0.0014), co-dominant (OR=0.66,  $P$ -value=8.5×10<sup>-6</sup>), and dominant models (OR=0.73,  $P$ -value=0.00073), while the association of the genetic variant and SBP, was significant under co-dominant model (coefficient=-0.74,  $P$ -value=0.047). The association of rs4762 on *AGT* and HTN was significant under the dominant (OR=1.24,  $P$ -value=0.01), co-dominant (OR=1.24,  $P$ -value=0.036), additive (OR=,  $P$ -value=0.016), and recessive models (OR=0.78,  $P$ -value=0.015).

### 3.4. Local Interaction findings

Local interaction within *AGT*, *TNXB*, and *ZBED9* for HTN, SBP, and DBP were investigated using the Mb-MDR algorithm, and only genetic variants on *ZBED9* had significant local interactions for HTN (Supplementary Table 4). Table 1 shows Mb-MDR results on HTN for significant combinations of genetic variants in a single locus of *ZBED9*. Two-locus antagonistic model of rs16901848 × rs450630 was significant for HTN (OR=0.72, *P*-value = $1.68 \times 10^{-5}$ ). The three-locus antagonistic model of rs16901848 × rs450630 × rs449074 was significant for HTN (OR= 0.73, *P*-value= $3.71 \times 10^{-5}$ ). The only synergistic model for HTN was rs450630 × rs449074 (OR=1.41, *P*-value= $2.81 \times 10^{-6}$ ).

Table 1  
The significant local interactions (SNP×SNP) on ZBED9 for hypertension

| GENE  | Interaction  | Model                                    | COEF  | Wald statistic | P-value               | FDR adjusted p-value | OR   |
|-------|--------------|--|-------|----------------|-----------------------|----------------------|------|
| ZBED9 | Antagonistic | rs450630(GA) ×rs16901848(AA)             | -0.32 | 18.52          | $1.68 \times 10^{-5}$ | <0.01                | 0.72 |
|       |              | rs450630(GA)×rs16901848(AA)×rs449074(CC) | -0.31 | 17.01          | $3.71 \times 10^{-5}$ | <0.01                | 0.73 |
|       | Synergistic  | rs450630(AA)×rs449074(TC)                | 0.35  | 21.94          | $2.81 \times 10^{-6}$ | <0.01                | 1.41 |

### 3.5. Epistasis analysis results

Table 2 presents the epistasis models of genetic variants in low LD ( $r^2 < 0.1$ ) on HTN, SBP, and DBP by Mb-MDR. Accordingly, rs450630 on *ZBED9* has significant interactions with different variants from *AGT* and *TNXB* in various combinations. Except for those containing rs2239689 on *TNXB*, the antagonistic combinations including, three-locus and four-locus, had heterozygous genotypes (Supplementary Table 3).

Table 2  
Epistatic (genexgene) Results of the Mb-MDR model

| Model Type         | Trait | Model  | Effect size*  | Wald statistic | P-value               | FDR adjusted P-value  |        |
|--------------------|-------|--|---|----------------|-----------------------|-----------------------|--------|
| Three locus models | HTN   | rs450630(GA)×rs12153855(CT)×rs2004776(CT)                | 0.73  | 17.09          | 3.57×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs12153855(CT)×rs4762(CT)                   | 0.73  | 15.7           | 7.43×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs12153855(AA)×rs2493134(TT)                | 0.6   | 16.4           | 5.13×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs12153855(CT)×rs11122576(GA)               | 0.66  | 25.76          | 3.87×10 <sup>-7</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2021783(CT)×rs2004776(GA)                 | 0.67  | 24.28          | 8.33×10 <sup>-7</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2021783(CT)×rs11122576(CT)                | 0.67  | 16.18          | 5.75×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2239689(CT)×rs2004776(CT)                 | 0.68  | 13.96          | 1.87×10 <sup>-4</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2239689(AA)×rs4762(AA)                    | 0.69  | 19.74          | 8.88×10 <sup>-6</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2269429(CT)×rs4762(TC)                    | 0.72  | 18.72          | 1.52×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2269429(CT)×rs2493134(TC)                 | 0.72  | 20.02          | 7.68×10 <sup>-6</sup> | <0.001                |        |
| Four locus models  | HTN   | rs450630(GA)×rs11122576(TC)×rs4762(GA)×rs12153855(CT)    | 0.69  | 19.24          | 1.15×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs11122576(CC)×rs4762(AA)×rs2239689(AA)     | 0.69  | 18.89          | 1.38×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs11122576(TC)×rs4762(GA)×rs2269429(CT)     | 0.69  | 18.08          | 2.12×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2004776(TT)×rs4762(AA)×rs2239689(AA)      | 0.6   | 22.55          | 2.05×10 <sup>6</sup>  | <0.001                |        |
|                    |       | rs450630(GA)×rs2004776(CT)×rs4762(GA)×rs2021783(CT)      | 0.68  | 22.03          | 2.69×10 <sup>6</sup>  | <0.001                |        |
|                    |       | rs450630(GA)×rs2004776(CT)×rs4762(GA)×rs2269429(CT)      | 0.68  | 17.99          | 2.22×10 <sup>5</sup>  | <0.001                |        |
|                    |       | rs450630(GA)×rs2493134(CT)×rs11122576(TC)×rs2269429(CT)  | 0.66  | 16.29          | 5.44×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs2493134(CT)×rs11122576(TC)×rs12153855(CT) | 0.68  | 15.4           | 8.72×10 <sup>-5</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs449074(TC)×rs4762(GA)×rs2021783(CT)       | 0.66  | 26.25          | 3×10 <sup>-7</sup>    | <0.001                |        |
|                    |       | rs450630(GA)×rs449074(TC)×rs4762(GA)×rs2269429(CT)       | 0.66  | 24.43          | 7.7×10 <sup>-7</sup>  | <0.001                |        |
|                    |       | rs450630(GA)×rs449074(TC)×rs2493134(CT)×rs12153855(CT)   | 0.67  | 16.04          | 6.22×10 <sup>-5</sup> | 0.02                  |        |
|                    |       | rs450630(GA)×rs449074(CC)×rs4762(AA)×rs2239689(AA)       | 0.67  | 21.53          | 3.49×10 <sup>-6</sup> | <0.001                |        |
|                    |       | rs450630(GA)×rs449074(TC)×rs2004776(CT)×rs12153855(CT)   | 0.71  | 18.59          | 1.62×10 <sup>5</sup>  | <0.001                |        |
|                    |       | rs450630(GA)×rs449074(TC)×rs2004776(CT)×rs2021783(CT)    | 0.72  | 19.67          | 9.20×10 <sup>6</sup>  | <0.001                |        |
|                    |       | rs450630(GA)×rs449074(TC)×rs11122576(TC)×rs12153855(CT)  | 0.72  | 17.29          | 3.20×10 <sup>-5</sup> | <0.001                |        |
|                    |       | SBP  | rs450630(GA)×rs2021783(CT)×rs2239689(GA)×rs2493134(CT)  | -5.49          | 16.4                  | 5.21×10 <sup>-5</sup> | <0.001 |
|                    |       | DBP  | rs450630(GA)×rs2239689(GA)×rs2493134(CT)×rs11122576(TC) | -2.95          | 22.84                 | 1.82×10 <sup>-6</sup> | <0.001 |
|                    |       |  | rs450630(GA)×rs2239689(GA)×rs2493134(CT)×rs2004776(CT)  | -2.11          | 20.77                 | 5.32×10 <sup>-6</sup> | <0.001 |

\*: Effect size for HTN is defined as odds ratio, and for SBP/DBP is defined as regression coefficient.

## 3.6. Entropy-based Interaction

Figure 2-A displays the entropy-based radial interactions of a combination of 11 biomarkers. The interaction effects are categorized by colours, as yellow indicates no-epistasis, green is for redundancy, blue is for the antagonistic association, and red is for the synergistic association. The joint effect sizes are written on connecting lines. After removing three variants with no effect on HTN, the interaction effects of 8 variants were depicted in Figure 2-B. As the main effects on the diagram shows, *ZBED9*: rs450630 (0.36%) and *AGT*: rs4762 (0.11%) had the largest antagonistic effect sizes on HTN.

## 3.7 Impact of genetic variants on ACE2 and TMPRSS2

Firstly, the sequence of *ZBED9* was analyzed. Figure 1 shows the disordered regions and disordered binding regions of this protein. rs450630 shows the ASP 686>GLU, the residue 686 of *ZBED9* belong to a region with the score ~0.5 for binding.

## 4. Discussion

The purpose of this study was to discover the interaction patterns involved between recently reported associated genes to unravel susceptible multi-locus in the regulation of blood pressure. In the first step, the integrated effect of *ZBED9*, *AGT*, and *TNXB* was assessed using a SKAT-O. Next, eleven variants were picked based on an LD threshold and relief algorithm. The independent effect of each variant was assessed by single-locus analysis, in which *ZBED9*: rs450630 and *AGT*: rs4762 were associated with hypertension. Finally, Mb-MDR and entropy-based methods were considered to draw the epistasis patterns between and within *AGT*, *ZBED9*, and *TNXB*. The Mb-MDR algorithm showed the significant synergistic and antagonistic effects of multi-locus local interactions for *ZBED9* and epistasis effects between *AGT*, *ZBED9*, and *TNXB*.

As a complex disease, high blood pressure stems from the interactions between multiple loci as well as environmental factors (24). Epistasis analysis supports joint effects of previous findings on the associations of *ZBED9* in GWAS and *AGT* and *TNXB* in linkage analysis with high blood pressure events in the Iranian population. Several studies have reported that interaction may unravel the main effects of genetic variants that are not associated in the single-locus analysis (25,26). Further, epistasis effects may explain genetic contribution on high blood pressure somewhat and disclose the mystery behind missing heritability (27).

Our results contribute to previous research on hypertension, where the protective effect of rs450630 (G>A) was reported in patients with hypertension. Results on independent effects of *ZBED9*: rs450630 showed that this variant exhibits a strong association with hypertension, and the results of epistasis and local interaction exacerbate this association, in which all the interaction combinations, whether with a homozygous or heterozygous genotype of other variants, were determined as antagonistic if *ZBED9*: rs450630 has heterozygous genotype. Hence, it appears that *ZBED9*: rs450630 has a deterministic role in the interaction patterns of the genetic variants on hypertension.

As mentioned above, Mb-MDR is a semiparametric machine learning dimension reduction algorithm. It is clear that Mb-MDR is a genotype-combination-based interaction testing method in which each p-locus model works with all combinations of genotype levels. This feature, along with the resampling-based approach, makes the Mb-MDR be less conservative than regression-based methods like the Haseman-Elston regression (28). In the Haseman-Elston regression method, the interaction term is treated as a multiplication of SNPs scores in the model, while some information is likely missed. Mb-MDR discovers significant SNP×SNP and gene×gene interactions with disease susceptibility in both case-control and quantitative association studies in unrelated individuals. Furthermore, it has been reported that it gains high power in the presence of genetic heterogeneity, genotyping error, missing data and low sample size(29).

A family-based genome-wide linkage study of blood pressure traits (SBP, DBP, and THN) was conducted via a two-level Haseman-Elston regression model on the TCGS dataset(9). In this study, in addition to the significant main effects of genetic variants on *ZBED9*, the epistasis between *AGT* and *TNXB* was also detected. However, our research uncovered some other significant interactions through checking for different patterns of genotypes on *ZBED9*, *AGT* and *TNXB*.

## Conclusion

We investigated the interplay between and within previously approved significant genes associated with high blood pressure traits (SBP, DBP, and THN) to determine the involved interaction patterns. To achieve this goal, following assessment of the aggregated effect

of the *ZBED9*, *AGT*, and *TNXB* genes, making the collection of relevant variants based on LD threshold and relifF algorithm, and testing single SNP effects, the local interactions and epistasis were tested via two methods; Mb-MDR, and entropy-based methods. Compared to previously utilized models, two-level Haseman-Elston regression, on the Iranian population, our strategy provided stronger evidence for the existence of interaction and epistasis affecting high blood pressure traits.

## Declarations

### Author Contributions

**M.A:** Conceptualization, Writing - Original Draft, Writing- Reviewing and Editing, Software, Formal analysis. **P.R:** Software, Formal analysis, Writing - Original Draft. **G.K and S.S:** Validation. **L.NHB and NA:** Visualization. **M.M:** Phenotype Validation. **H.L:** Bioinformatic analysis. **F.A:** Supervision. **M.SD:** Supervision, Writing- Reviewing, and Editing. All authors read and approved the final manuscript.

### Corresponding author

Correspondence to **Maryam S Daneshpour**.

### Funding

All parts of this research work, design of the study, data collection, analysis, interpretation of data, and writing the manuscript, were funded by the Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. The funding body played no role in publication costs and also recognizes the scientific support of deCODE (Reykjavik, Iceland).

### Acknowledgements

The authors would like to express their gratitude to the staff and participants in the TCGS project. Also, special thanks for the scientific and financial support of the deCODE genetic company (Reykjavik, Iceland).

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Availability of data and materials

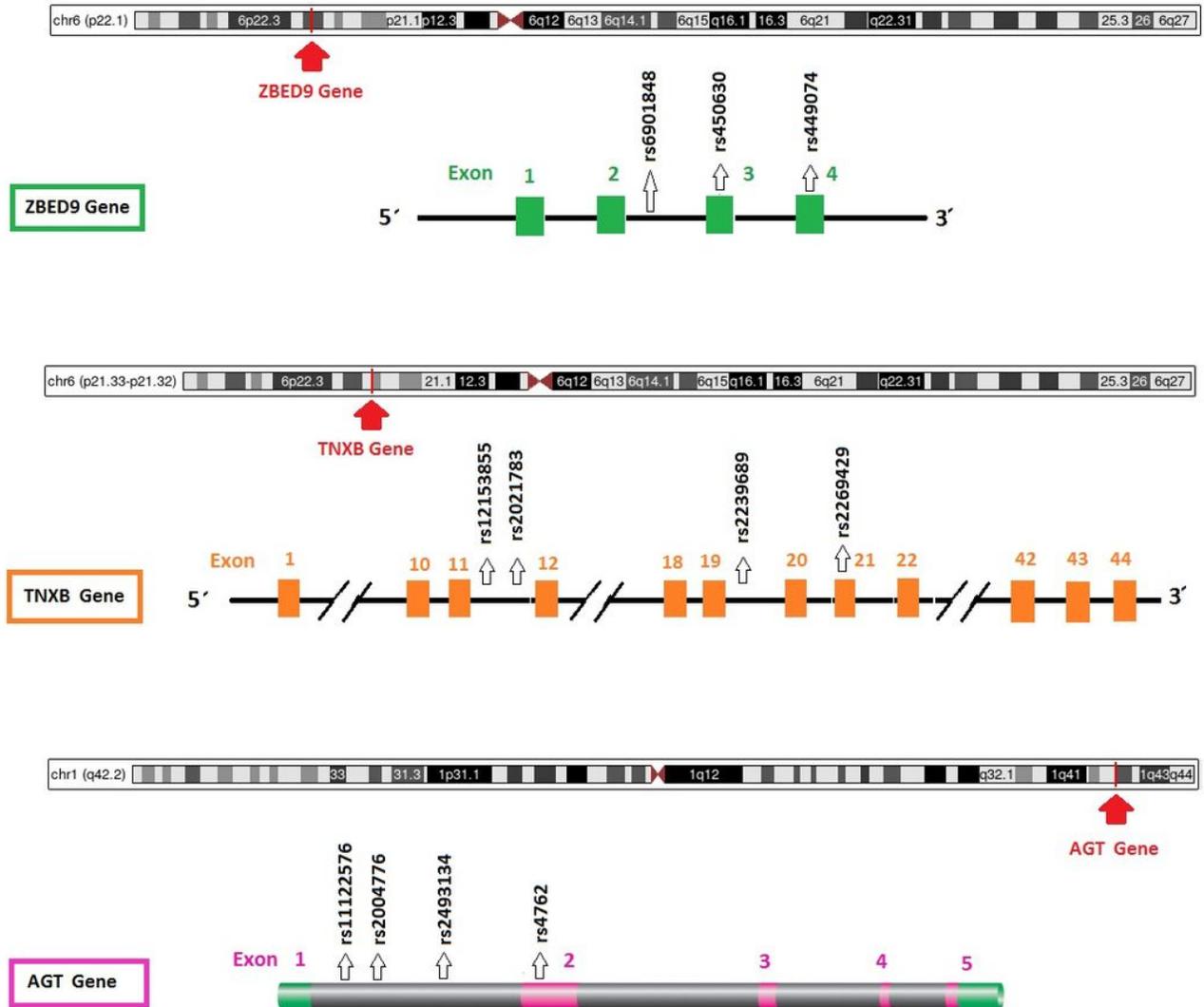
The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

## References

1. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115mmHg, 1990-2015. *JAMA - Journal of the American Medical Association* [Internet]. 2017 Jan;317(2):165–82. Available from: <https://pubmed.ncbi.nlm.nih.gov/28097354/>
2. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nature Reviews Nephrology* [Internet]. 2020;16(4):223–37. Available from: <http://dx.doi.org/10.1038/s41581-019-0244-2>
3. Mills KT, Bundy JD, Kelly TN, Reed JE, Kearney PM, Reynolds K, et al. Global disparities of hypertension prevalence and control. *Circulation* [Internet]. 2016 Aug;134(6):441–50. Available from: </pmc/articles/PMC4979614/>
4. Ibrahim MM, Damasceno A. Hypertension in developing countries. *The Lancet* [Internet]. 2012;380(9841):611–9. Available from: [http://dx.doi.org/10.1016/S0140-6736\(12\)60861-7](http://dx.doi.org/10.1016/S0140-6736(12)60861-7)
5. Kolifarhood G, Daneshpour M, Hadaegh F, Sabour S, Saadati HM, Haghdoost AA, et al. Heritability of blood pressure traits in diverse populations: a systematic review and meta-analysis. *Journal of human hypertension*. 2019;33(11):775–85.
6. Kolifarhood G, Daneshpour MS, Khayat BS, Saadati HM, Guity K, Khosravi N, et al. Generality of genomic findings on blood pressure traits and its usefulness in precision medicine in diverse populations: A systematic review. *Clinical Genetics*. 2019;96(1).

7. Bijani M, Parvizi S, Dehghan A, Sedigh-Rahimabadi M, Rostami-chijan M, Kazemi M, et al. Investigating the prevalence of hypertension and its associated risk factors in a population-based study: Fasa PERSIAN COHORT data. *BMC Cardiovascular Disorders*. 2020 Dec;20(1).
8. Khosravi A, Emamian MH, Shariati M, Hashemi H, Fotouhi A. The prevalence of pre-hypertension and hypertension in an Iranian urban population. *High Blood Pressure and Cardiovascular Prevention*. 2014;21(2):127–35.
9. Kolifarhood G, Daneshpour MS, Zahedi AS, Khosravi N, Sedaghati-Khayat B, Guity K, et al. Familial genetic and environmental risk profile and high blood pressure event: a prospective cohort of cardio-metabolic and genetic study. *Blood Pressure [Internet]*. 2021;30(3):196–204. Available from: <https://www.tandfonline.com/doi/abs/10.1080/08037051.2021.1903807>
10. Kolifarhood G, Sabour S, Akbarzadeh M, Sedaghati-khayat B, Guity K, Dehkordi SR, et al. Genome-Wide Association Study on Blood Pressure Traits in the Iranian Population Suggests ZBED9 as A New Locus for Hypertension. *Research Square*. 2020;1–20.
11. Daneshpour MS, Hedayati M, Sedaghati-Khayat B, Guity K, Zarkesh M, Akbarzadeh M, et al. Genetic identification for non-communicable disease: Findings from 20 years of the Tehran Lipid and Glucose Study. *International journal of endocrinology and metabolism*. 2018;16(4 Suppl).
12. Daneshpour MS, Fallah M-S, Sedaghati-Khayat B, Guity K, Khalili D, Hedayati M, et al. Rationale and Design of a Genetic Study on Cardiometabolic Risk Factors: Protocol for the Tehran Cardiometabolic Genetic Study (TCGS). *JMIR Research Protocols [Internet]*. 2017 Feb;6(2):e28. Available from: <https://pubmed.ncbi.nlm.nih.gov/28232301/>
13. Akbarzadeh M, Riahi P, Ramezankhani A, Dehkordi SR, Roudbar MA, Zarkesh M, et al. Parental transmission plays the major role in the high aggregation of type 2 Diabetes in Iranian families: Tehran Lipid and Glucose Study (TLGS). *Canadian Journal of Diabetes*. 2021 Jun 1;
14. Azizi F, Rahmani M, Emami H, Madjid M. Tehran lipid and glucose study: Rationale and design. *CVD Prevention [Internet]*. 2000 [cited 2020 Dec 30];3(3):242–7. Available from: [http://ijem.sbmu.ac.ir/browse.php?a\\_code=A-10-6-10&slc\\_lang=en&sid=1&sw=Gene](http://ijem.sbmu.ac.ir/browse.php?a_code=A-10-6-10&slc_lang=en&sid=1&sw=Gene)
15. Kolifarhood G, Daneshpour MS, Zahedi AS, Khosravi N, Sedaghati-Khayat B, Guity K, et al. Familial genetic and environmental risk profile and high blood pressure event: a prospective cohort of cardio-metabolic and genetic study. *Blood Pressure*. 2021;30(3):196–204.
16. <https://www.researchprotocols.org/2017/2/e28>
17. Akbarzadeh M, Dehkordi SR, Roudbar MA, Sargolzaei M, Guity K, Sedaghati-Khayat B, et al. GWAS findings improved genomic prediction accuracy of lipid profile traits: Tehran Cardiometabolic Genetic Study. *Scientific reports*. 2021;11(1):1–9.
18. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics*. 2012;13(4):762–75.
19. Lee SS. SKAT Package. 2014;1–16.
20. Kira K, Rendell LA. A Practical Approach to Feature Selection. *Machine Learning Proceedings 1992*. 1992 Jan 1;249–56.
21. Wang X. Firth logistic regression for rare variant association tests. *Frontiers in Genetics [Internet]*. 2014 Jun;5(JUN):187. Available from: [www.frontiersin.org](http://www.frontiersin.org)
22. Calle ML, Gales VU, Riera NM i, Steen K van. MB-MDR: model-based multifactor dimensionality reduction for detecting interactions in high-dimensional genomic data. 2008; Available from: <http://dspace.uvic.ca/handle/10854/408>
23. Kwon MS, Park M, Park T. IGENT: Efficient entropy based algorithm for genome-wide gene-gene interaction analysis. *BMC Medical Genomics*. 2014;7(SUPPL.1):1–11.
24. Neale BM, Ferreira MAR, Medland SE, Posthuma D, Taylor Q. *Statistical Genetics: Gene Mapping through Linkage and Association [Internet]*. 2008. Available from: <http://www.genemapping.org>
25. Lou Y, Liu J, Huang Y, Liu J, Wang Z, Liu Y, et al. A46G and C79G polymorphisms in the B2-adrenergic receptor gene (ADRB2) and essential hypertension risk: A meta-analysis. *Hypertension Research*. 2010;33(11):1114–23.
26. Harrap SB. Blood pressure genetics: time to focus [Internet]. Vol. 3, *Journal of the American Society of Hypertension*. *J Am Soc Hypertens*; 2009. p. 231–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/20409965/>
27. Kupper N, Willemsen G, Riese H, Posthuma D, Boomsma DI, de Geus EJC. Heritability of daytime ambulatory blood pressure in an extended twin design. *Hypertension*. 2005 Jan;45(1):80–5.
28. Wang T, Elston RC. Two-level Haseman-Elston regression for general pedigree data analysis. *Genetic Epidemiology [Internet]*. 2005 Jul 1 [cited 2021 Aug 24];29(1):12–22. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/gepi.20075>

## Figures



**Figure 1**

This diagram depicts the three genes ZBED9, TNXB, and AGT, as well as the locations of 11 significant polymorphisms in these genes on chromosomes 1 and 6. As illustrated in the figure, the ZBED9 gene contains two polymorphisms in exons 3 and 4, and one polymorphism in intron region 2. One polymorphism is located in exon 21 and the other three are located in the introns of the TNXB gene. Additionally, the AGT gene contains a variant in exon 2 and others in the intronic regions, and the interaction of these polymorphisms results in a significant association with hypertension.

**Figure 2**

Independent effects of each variant are written on each node. Interaction effects are defined by colors; Yellow is indicative of no-epistasis, green is for redundancy, blue is for the antagonistic association, and red is for the synergistic association. The strength of the

correlation is written on connecting lines.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable3.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [Supplementaryinformation.docx](#)