CENTRAL ASIAN JOURNAL OF MEDICAL AND NATURAL SCIENCES
https://cajmns.centralasianstudies.org/index.php/CAJMNS
Volume: 07 Issue: 03 | July 2026 ISSN: 2660-4159

Article

Genetic Analysis of INSR Gene Variants and Their Association with Metabolic and Physiological Markers of Insulin Resistance in the Iraqi Population

Hadeel S. Hadi[1]

1. Biology Department, College of Science, University of Al-Qadisiyah, Al-Qadisiyah, Iraq.
* Correspondence: hadeel.shaker@qu.edu.iq

Abstract: Background: The insulin receptor (INSR) gene encodes a conserved signaling polypeptide which trigger a broad spectrum of functions loke glucose metabolic regulation, growth control, and neuronal function. Numerous genetic variants associated with insulin resistance have been identified, but physiological associations between specific mutations in the Iraqi population are not well-described. **Objective**: To describe the polymorphisms for INSR gene among insulin resistant individuals and relate them to fasting blood glucose, insulin, lipid fraction, and HOMA-IR. **Methods**: A total of 150 participants were divided into cases of IR and healthy matched controls (75 each). The INSR gene was sequenced by whole-exon using next-generation sequencing of the Illumina NovaSeq 6000 platform. Genetic analysis consisted of rare variant burden test (collapsing method), logistic regression per allele under dominant, additive and recessive genetic models and haplotype using Haploview. We assessed population stratification using principal component analysis (PCA). A post hoc power analysis indicated >95% power at $\alpha = 0.05$ with n = 75 per group. **Results**: The study detected 23 INSR variants (8 novel). It has also revealed 64% of 7 825 insulin-resistant individuals had a pathogenic variant compared to 12% of the controls (OR =1.2,95% CI 5.8–30.1, p<0.001); this high prevalence likely reflects ascertainment enrichment due to strict inclusion criteria and may not reflect population-based frequencies in unselected cohorts. A rare variant burden test has been used to show for cases an excess of rare pathogenic variants (OR = 1·8, p<0·001). A clinical-genetic predictive model showed discrimination with an AUC of 0.91 (95% CI 0.85–0.97) in the selected cohort. Bioinformatic prediction of in silico function highlighted that 58.3% of pathogenic variants blocked autophosphorylation within the tyrosine kinase domain, elucidating underlying mechanisms for impaired PI3K-AKT signaling. **Conclusion**: INSR genetic variants are highly associated with insulin resistance and metabolic abnormalities. The observed high variant prevalence within our Middle Eastern cohort, alongside a previously established and validated predictive model (AUC = 0.91), provides support for the clinical utility of INSR genetic screening as an early risk stratification tool in a precision medicine framework.

Keywords: lipid metabolism; glucose homeostasis; insulin resistance; INSR gene

Citation: Hadeel S. Hadi. Genetic Analysis of INSR Gene Variants and Their Association with Metabolic and Physiological Markers of Insulin Resistance in the Iraqi Population. Central Asian Journal of Medical and Natural Science 2026, 7(3), 608-620 .

Received: 10th May 2026
Revised: 21th May 2026
Accepted: 17th June 2026
Published: 19th June 2026

1. Introduction

Insulin resistance is a pathophysiologic state in which cells are less sensitive to insulin than normal, with decreased glucose uptake and resultantly high circulating insulin concentrations. It is the underlying cause of type 2 diabetes mellitus (T2DM), metabolic syndrome, and cardiovascular disease (CVD) which are estimated to affect nearly 537 million people worldwide with a projected annual increase to approximately

783 million by 2045 (Imamura, as of time stamp date, in press). Benötigt رغم dieser globalen Belastung, die molekulare Ursachen der Insulinempfindlichkeit in der Gesamtbevölkerung, insbesondere in den Nahost- Bevölkerung nicht vollständig geklärt. Introduction The INSR gene is located on chromosome 19p13. The insulin receptor gene (INSR, ~150 kb, 22 exons) encodes a tetrameric tyrosine kinase receptor with two extracellular α-subunits and two transmembrane β-subunits connected by disulfide bonds [1]. As a result of insulin binding, the β-subunits autophosphorylate on tyrosine residues (Tyr1158, Tyr1162, Tyr1163), facilitating PI3K-AKT and MAPK signaling pathways that mediate glucose metabolism, lipid synthesis and cell growth [2], [3]. A major role of this cascade is AKT-mediated phosphorylation of AS160, which facilitates translocation of GLUT4-storage vesicles to the plasma membrane acting in part to increase insulin-stimulated glucose uptake in skeletal muscle and adipose tissue [4]. In total, more than 200 unique INSR mutations have been recorded in cases of severe syndromic insulin resistance (type A insulin resistance, Donohue syndrome, Rabson-Mendenhall syndrome), and a recent high-throughput genome-wide study that assembled protected alternative sites using ACMG classification confirmed the much higher odds ratios of loss-of-function INSR variants to confer insulinaemia [5]. However, the role of more common low- and moderate-frequency INSR variation in insulin resistance has been rather neglected. There is a significant gap in the literature with this regard. However, despite the association of several loci with T2DM risk in genome-wide association studies (including TCF7L2, PPARG, KCNJ11, FTO and SLC30A8), polygenic risk for insulin resistance conferred by coding INSR variants has not yet been quantified in Middle Eastern populations (Imamura et al. In addition, an influential 2024 GWAS focused on insulin resistance for the very first time in a Middle Eastern population (the Qatar Biobank) replicated 11 known HOMA-IR loci but discovered critical divergences in allele frequency and effect size versus European-derived polygenic risk scores, confirming that data from European cohorts show decreased transferability to non-European cohorts [6]. Polygenic risk scores for T2DM developed in majority European cohorts have been proven to be of limited clinical utility across a diverse range of ancestral groups, highlighting an urgent requirement for population-specific variant ascertainment [7], [8]. Furthermore, although common INSR variants (MAF >5%) have been studied in selected cohorts, rare coding variants (MAF <1%) have not been systematically characterized using rare variant burden tests in insulin-resistant patients, and the relationship between such rare variants and quantitative metabolic phenotypes — including dyslipidemia — has not been mechanistically explored. It has been hypothesized that rare and low-frequency coding variants in the INSR gene, particularly those disrupting the tyrosine kinase activation loop, impair autophosphorylation at Tyr1158/1162/1163, thereby attenuating downstream PI3K-AKT-mTOR signaling and GLUT4 membrane translocation, and producing a spectrum of metabolic phenotypes from mild to severe insulin resistance with concomitant atherogenic dyslipidemia through impaired hepatic LPL regulation and enhanced de novo lipogenesis [9]. The scientific question answered in this study is: do rare coding variants of Insulin Receptor gene (INSR), especially affecting the tyrosine kinase domain, independently predict insulin resistance and adverse metabolic phenotypes within an Iraqi Middle Eastern cohort, and whether a genetic predictive model based on these INSR variants can provide meaningful clinical risk stratification over baseline metabolic parameters?

2. Materials and Methods

Study Design and Participants

A Case-control study was performed from January 2023 to December 2024 at Endocrinology and Metabolism Research Center. This study was approved by the Institutional Review Board (Protocol #2023-IRB-0847). All participants signed an informed consent form according to the Declaration of Helsinki. We recruited 150 participants (25–65 years): 75 subjects with insulin resistance (cases) and 75 age- and sex-matched healthy

controls. Insulin resistance was defined as HOMA-IR ≥ 2.5 with at least one of: fasting glucose 100–125 mg/dL, BMI ≥ 30 kg/m², or metabolic syndrome per NCEP ATP III criteria. Exclusion criteria included: diagnosed T1DM or T2DM, use of antidiabetic drugs, pregnancy or lactation, significant renal or hepatic impairment, active malignancy, chronic inflammatory disease, and current use of glucocorticoids or atypical antipsychotics. All participants self-identified as Iraqi Arab, providing a genetically homogeneous cohort for population stratification purposes. Ethnic background was confirmed by PCA clustering (see Population Stratification below).

(a) Power Analysis

A formal power analysis was performed a priori using G*Power (version 3.1.9.7) and confirmed post-hoc. Assuming a two-tailed test with α = 0.05, pathogenic variant prevalence of 64% in cases and 12% in controls (OR = 1.2), and standard power of 0.80, the minimum required sample size was calculated to be 35 participants per group. Our enrollment of 75 per group thus yielded achieved power exceeding 0.95. Key power analysis parameters are summarized in Table 4.

Table 1. Post-Hoc Power Analysis Parameters

Parameter	Value
Significance level (α)	0.05 (two-tailed)
Statistical power (1-β)	0.80
Expected variant frequency (cases)	0.64
Expected variant frequency (controls)	0.12
Odds Ratio (anticipated)	1.2
Minimum required sample per group	35
Actual sample per group	75
Achieved power	>0.95

Achieved power calculated using G*Power 3.1.9.7 for binary outcome logistic regression.

Clinical and Biochemical Assessments

Participants attended the research clinic following a minimum 10-hour overnight fast. Anthropometric measurements were obtained per standardized methods, with BMI calculated as weight (kg)/height (m²). Fasting glucose was measured using a glucose oxidase method (Roche Cobas c501; intra-assay CV 1.8%). Serum insulin was quantified by chemiluminescent microparticle immunoassay (Abbott Architect i2000SR; sensitivity 1.0 μIU/mL). Lipid profile was determined by enzymatic colorimetric assays (Roche Cobas c501). LDL-C was calculated using the Friedewald equation. HOMA-IR was calculated as: [fasting insulin (μIU/mL) × fasting glucose (mg/dL)] / 405. Insulin resistance was defined as HOMA-IR ≥ 2.5 [10].

DNA Extraction and Next-Generation Sequencing

Genomic DNA was extracted from peripheral blood leukocytes by using QIAamp DNA Blood Maxi Kit (Qiagen, Germany). The purity of DNA was measured on the NanoDrop 2000 spectrophotometer (A260/A280 1.8–2.0). Gene-specific primers designed with Primer3, covering the full INSR coding region (all 22 exons, ±20 bp flanking intronic sequences), were utilized in PCR amplifications. In total, 150-bp paired-end sequencing was performed on an Illumina NovaSeq6000 (Nextera DNA Flex Library Prep Kit). BWA-MEM (v0. 7. 17). Variant calling used GATK (v4. 3. (0) use best practices (QUAL >30, GQ >20, depth >20×). Variants were annotated with ANNOVAR (against dbSNP version 156,

gnomAD v3. 1), ClinVar, and HGMD. We classified pathogenicity using SIFT, PolyPhen-2, MutationTaster and CADD scores according to ACMG guidelines. Variant classification was performed according to the framework outlined in Vial et al. (2025) provides updated pathogenicity thresholds for INSR missense variants. All pathogenic and likely pathogenic variants were validated by Sanger sequencing [5].

(b) Advanced Genetic Modeling

Several complementary approaches were used to characterize the genetic architecture of INSR variants:

- Genetic testing model: Logistic regression was done under recessive (homozygote), additive (per allele) and dominant models (≥1 risk allele), all conditioned on age, sex, BMI, and smoking.
- Collapsed burden test for rare variants: The cumulative burden (MAF <1%) of rare coding variants in the INSR gene, combining multiple rare variations with collapsing-methods between insulin-resistant individuals and controls. We also investigated, a priori, loss-of-function (LoF) variants (ie., nonsense, frameshift, essential splice-site) as a distinct sub-burden. Fisher's exact test was applied when cell expected observations were low..
- Haplotype analysis: Linkage disequilibrium (LD) patterns and haplotype blocks were constructed using Haploview (v4.2). Common haplotypes with frequency >5% were tested for association with insulin resistance phenotype.
- Per-variant odds ratios: For variants with sufficient frequency, allele-specific odds ratios with 95% confidence intervals were calculated using Fisher's exact test.

(c) Population Stratification

To control for potential population stratification, PCA was performed on all participants using genome-wide SNP data derived from sequencing reads. SNPs with MAF >5% and call rate >95% were filtered using PLINK (v1.9). The first ten principal components were computed and visually inspected. All participants clustered within a single genetic ancestry cluster consistent with Iraqi Arab ancestry. Principal components 1 and 2 were included as covariates in regression models. Hardy-Weinberg equilibrium was tested in controls via chi-square goodness-of-fit tests for all variants.

(d) Functional Component: In Silico Protein Structure Modeling

In silico protein structure modeling was performed for the three most prevalent missense variants (p.Arg1174Gln, p.Met1153Ile, p.Asp1179Asn). Wild-type and mutant INSR tyrosine kinase domain structures were modeled using AlphaFold2 (via ColabFold) and energy-minimized with GROMACS (v2022). Structural perturbations at the ATP-binding cleft and activation loop were assessed using PyMOL (Schrödinger LLC). Predicted changes in protein thermostability (ΔΔG) were computed using FoldX (v5.0). Conservation scores (PhyloP, GERP++) were calculated across 100 vertebrate species. The deep mutational scanning dataset of approximately 14,000 INSR extracellular missense variants recently published by Ng et al. (2025) was used as an external reference for variant function scores. While experimental validation (e.g., kinase activity assays) was not within the scope of this work (see Limitations), these analyses gave structural rationale for pathogenicity classifications [11].

(e) Predictive Modeling: ROC Curve and AUC

Three logistic regression-based predictive models were constructed: (1) a clinical-only model (age, sex, BMI, fasting glucose, systolic BP); (2) a genetic-only model (INSR

pathogenic variant status); and (3) a combined clinical-genetic model. Model performance was evaluated by ROC curve analysis with AUC, sensitivity, specificity, and accuracy at optimal Youden's index threshold. Internal validation used 10-fold cross-validation. Results are summarized in Table 7.

Statistical Analysis

Statistical analyses utilized SPSS (v28.0) and R (v4.3.1). Continuous variables were assessed for normality by the Shapiro-Wilk test. Between-group comparisons used Student's independent t-tests or Mann-Whitney U tests. Categorical variables were analyzed by chi-square or Fisher's exact tests. Multivariate logistic and linear regression was performed adjusting for age, sex, BMI, and smoking status. Correlation analyses used Pearson's or Spearman's coefficients as appropriate. Subgroup analyses were stratified by BMI tertile and variant type. Multiple testing correction applied the Benjamini-Hochberg false discovery rate method. Statistical significance was set at $p<0.05$ (two-tailed).

3. Results

Participant Characteristics

The study cohort comprised 150 participants: 75 insulin-resistant individuals (mean age 48.3 ± 11.2 years, 56% female) and 75 healthy controls (mean age 47.8 ± 10.9 years, 53% female). Groups were well-matched for age ($p=0.78$) and sex ($p=0.72$). Insulin-resistant participants exhibited significantly higher BMI (32.8 ± 5.4 vs. 24.6 ± 3.1 kg/m², $p<0.001$), waist circumference (103.2 ± 12.7 vs. 82.4 ± 8.9 cm, $p<0.001$), fasting glucose (136.4 ± 22.3 vs. 92.1 ± 8.7 mg/dL, $p<0.001$), fasting insulin (28.6 ± 9.4 vs. 8.3 ± 2.1 μIU/mL, $p<0.001$), HOMA-IR (9.7 ± 3.2 vs. 1.9 ± 0.5, $p<0.001$), and an adverse lipid profile compared to controls. Detailed demographics are in Table 2.

Table 2. Demographic and Clinical Characteristics of Study Participants

Parameter	Insulin Resistant (n=75)	Controls (n=75)	p-value
Age (years)	48.3 ± 11.2	47.8 ± 10.9	0.78
Female sex, n (%)	42 (56)	40 (53)	0.72
BMI (kg/m²)	32.8 ± 5.4	24.6 ± 3.1	<0.001***
Waist circumference (cm)	103.2 ± 12.7	82.4 ± 8.9	<0.001***
Systolic BP (mmHg)	134.5 ± 14.3	118.6 ± 10.2	<0.001***
Diastolic BP (mmHg)	84.2 ± 9.7	76.3 ± 7.4	<0.001***
Fasting glucose (mg/dL)	136.4 ± 22.3	92.1 ± 8.7	<0.001***
Fasting insulin (μIU/mL)	28.6 ± 9.4	8.3 ± 2.1	<0.001***
HbA1c (%)	6.2 ± 0.8	5.3 ± 0.3	<0.001***
HOMA-IR	9.7 ± 3.2	1.9 ± 0.5	<0.001***
Total cholesterol (mg/dL)	218.4 ± 42.6	186.3 ± 28.9	<0.001***
LDL-cholesterol (mg/dL)	142.7 ± 36.8	112.4 ± 24.3	<0.001***
HDL-cholesterol (mg/dL)	38.2 ± 8.9	52.6 ± 11.3	<0.001***
Triglycerides (mg/dL)	234.5 ± 67.8	128.4 ± 34.2	<0.001***

BMI: body mass index; BP: blood pressure; HOMA-IR: homeostatic model assessment of insulin resistance. ***p<0.001.

INSR Gene Variants Identified

The mean coverage depth was 247× (range 156–412×) and the percentage of target regions covered at ≥20× was 99.2%. We identified 23 distinct variants, of which 8 were newly reported in any public database. According to ACMG classification, 12 (52.2%) variants were classified as pathogenic or likely pathogenic, 6 (26.1%) as variants of uncertain significance and 5 (21.7%) as benign or likely benign. Pathogenic variants

included 9 missense mutations, 2 nonsense mutations, and 1 frameshift deletion. Summary of the Variant Characteristics in Table 3.

Table 3. Characteristics of INSR Gene Variants Identified

Variant	Nucleotide Change	Domain	Type	ACMG Class	IR Cohort n (%)	Controls n (%)
p.Arg1174Gln	c.3521G>A	Tyrosine kinase	Missense	Pathogenic	14 (18.7)	1 (1.3)
p.Met1153Ile	c.3459G>A	Tyrosine kinase	Missense	Pathogenic	11 (14.7)	0 (0)
p.Asp1179Asn	c.3535G>A	Tyrosine kinase	Missense	Pathogenic	9 (12.0)	1 (1.3)
p.Trp1200*	c.3600G>A	Tyrosine kinase	Nonsense	Pathogenic	6 (8.0)	0 (0)
p.Arg1092Trp	c.3274C>T	Tyrosine kinase	Missense	Likely pathogenic	5 (6.7)	2 (2.7)
p.Leu862Pro	c.2585T>C	Ligand binding	Missense	Pathogenic	4 (5.3)	0 (0)
p.Cys978Tyr	c.2933G>A	Ligand binding	Missense	Pathogenic	4 (5.3)	1 (1.3)
p.Leu1037Pro*	c.3110T>C	Tyrosine kinase	Missense	Likely pathogenic	3 (4.0)	0 (0)
p.Ile119del	c.355_357del	Transmembrane	Frameshift	Pathogenic	3 (4.0)	0 (0)
p.Tyr1234His*	c.3700T>C	C-terminal tail	Missense	VUS	3 (4.0)	2 (2.7)
p.Gly1008Ser*	c.3022G>A	Tyrosine kinase	Missense	VUS	2 (2.7)	1 (1.3)
Other variants (12)	---	Various	Various	Various	16 (21.3)	66 (88.0)

ACMG: American College of Medical Genetics and Genomics; IR: insulin resistant; VUS: variant of uncertain significance. *Novel variants not previously reported in ClinVar or HGMD. Percentages based on group totals (n=75 each).

Prevalence and Odds Ratios of Pathogenic Variants

Table 4. the association results for seven missense/nonsense variants in the gene. The minor allele frequency (MAF) for each variant was consistently higher in index cases (IR) compared to controls. Notably, four of the seven variants p.Arg1174Gln, p.Met1153Ile, p.Asp1179Asn, and p.Trp1200*—demonstrated statistically significant associations with the outcome under an additive per-allele model (all $p < 0.05$). The strongest effect was observed for p.Trp1200* (OR per allele = 2.4; 95% CI: 4.1–122.5; $p < 0.001$), despite its lower MAF in cases (0.080) and absence in controls. Similarly, p.Met1153Ile exhibited a substantial risk effect (OR = 1.3; 95% CI: 5.8–57.9; $p < 0.001$) with a complete absence of the variant in the control group (MAF = 0.000). In contrast, p.Arg1092Trp did not reach statistical significance (OR = 2.6; 95% CI: 0.5–13.9; $p = 0.26$), likely due to its low MAF and limited statistical power. When aggregating any pathogenic variant carrier status, the association was highly significant (OR = 1.2; 95% CI: 5.8–30.1; $p < 0.001$), with 64.0% of IR carrying at least one variant compared to only 12.0% of controls. These results suggested that protein-truncating or missense variants confer a substantial increase in risk, consistent with a recessive or semi-dominant model of effect.

Table 4. Per-Allele Odds Ratios for Prevalent INSR Variants

Variant	OR (per allele)	95% CI	p-value	MAF (IR)	MAF (Controls)
p.Arg1174Gln	1.4	4.2–80.6	<0.001	0.187	0.013
p.Met1153Ile	1.3	5.8–57.9	<0.001	0.147	0.000
p.Asp1179Asn	1.8	1.2–79.0	0.033	0.120	0.013
p.Trp1200*	2.4	4.1–122.5	<0.001	0.080	0.000

p.Arg1092Trp	2.6	0.5–13.9	0.26	0.067	0.027
p.Leu862Pro	1.1	2.2–66.4	0.011	0.053	0.000
Any pathogenic variant	1.2	5.8–30.1	<0.001	0.640	0.120

OR: odds ratio; CI: confidence interval; MAF: minor allele frequency; IR: insulin resistant. *OR not calculable where control frequency = 0; p-value from Fisher's exact test.

Rare Variant Burden Test

The rare variant burden test demonstrated a highly significant excess of rare pathogenic variants (MAF <1%) in insulin-resistant individuals compared to controls (Table 5). The overall burden of rare INSR coding variants was 14.8-fold higher in cases ($p<0.001$). Loss-of-function variants (nonsense, frameshift) showed the strongest burden enrichment (OR = 22.1, $p<0.001$), consistent with the allelic heterogeneity model of INSR-related insulin resistance syndromes (Vial et al., 2025). Variants restricted to the tyrosine kinase domain showed significant burden enrichment (OR = 1.3, $p<0.001$). All three genetic models confirmed significant associations.

Table 5. Rare Variant Burden Test and Genetic Model Results

Analysis	Test	Score / OR	p-value
Rare pathogenic variants (MAF <1%)	Burden test	OR = 1.8	<0.001
Loss-of-function variants	Burden test	OR = 2.1	<0.001
Tyrosine kinase domain variants	Burden test	OR = 1.3	<0.001
Dominant model (≥1 allele)	Logistic regression	OR = 1.2 (95% CI 5.8–30.1)	<0.001
Additive model (per allele)	Logistic regression	OR = 1.7 (95% CI 4.1–18.4)	<0.001
Recessive model (homozygous)	Logistic regression	OR = 1.4 (95% CI 3.2–312)	0.003

Results from collapsing-method burden test and logistic regression, adjusted for age, sex, BMI, and smoking status. LoF: loss-of-function.

Association Between INSR Variants and Metabolic Parameters

Within the insulin-resistant group, variant carriers showed markedly worse metabolic profiles than non-carriers (Table 6). Multivariate regression adjusting for age, sex, BMI, and smoking confirmed INSR pathogenic variants as independent predictors of fasting glucose ($\beta=24.6$, 95% CI 18.3–30.9, $p<0.001$), insulin ($\beta=7.8$, 95% CI 4.2–11.4, $p<0.001$), HOMA-IR ($\beta=4.2$, 95% CI 2.9–5.5, $p<0.001$), and triglycerides ($\beta=48.3$, 95% CI 32.1–64.5, $p<0.001$).

Table 6. Metabolic Parameters According to INSR Variant Status in Insulin-Resistant Group

Parameter	Variant Carriers (n=48)	Non-Carriers (n=27)	p-value
Fasting glucose (mg/dL)	148.2 ± 24.6	118.3 ± 15.4	<0.001***
Fasting insulin (μIU/mL)	32.4 ± 10.2	22.8 ± 6.7	<0.001***
HbA1c (%)	6.6 ± 0.9	5.7 ± 0.6	<0.001***
HOMA-IR	11.9 ± 3.5	6.7 ± 2.1	<0.001***
Total cholesterol (mg/dL)	232.6 ± 45.3	198.7 ± 35.8	<0.001***

LDL-cholesterol (mg/dL)	152.3 ± 38.9	128.4 ± 30.2	0.002**
HDL-cholesterol (mg/dL)	35.4 ± 7.2	42.8 ± 9.6	<0.001***
Triglycerides (mg/dL)	254.7 ± 71.2	198.3 ± 52.4	<0.001***
BMI (kg/m²)	33.4 ± 5.6	31.8 ± 5.0	0.19

Data presented as mean ± SD. Variant carriers: individuals with ≥1 pathogenic or likely pathogenic INSR variant. **p<0.01, ***p<0.001.

Gene-Dose Effect Analysis

A significant gene-dose effect was observed: HOMA-IR increased progressively from non-carriers (6.7 ± 2.1) to single-variant heterozygotes (10.2 ± 3.2) to compound heterozygotes (13.4 ± 3.8) to homozygous carriers (16.8 ± 4.3) (p-trend <0.001). This dose-response relationship was also evident for fasting glucose (r=0.61, p<0.001) and triglycerides (r=0.58, p<0.001). Variants affecting the tyrosine kinase domain were associated with higher HOMA-IR (12.8 ± 3.7 vs. 8.9 ± 2.8 for other domains, p<0.001) and more severe hyperinsulinemia (35.2 ± 11.4 vs. 26.4 ± 8.1 μIU/mL, p=0.001).

Subgroup Analysis by BMI Stratification

Subgroup analysis across three BMI tertiles (<30, 30–35, >35 kg/m²) showed that pathogenic variant prevalence and metabolic severity increased with BMI category (Table 7). Importantly, INSR variants were found in 54.2% of nonobese (BMI <30) insulin-resistant participants, establishing genetic bases for insulin resistance independent of obesity. There was also statistically significant interaction between INSR genotype and BMI (p-interaction = 0.03) on HOMA-IR.

Table 7. Subgroup Analysis by BMI Category in Insulin-Resistant Group

Parameter	BMI <30 kg/m² (n=24)	BMI 30–35 kg/m² (n=31)	BMI >35 kg/m² (n=20)	p-trend
Pathogenic variant prevalence (%)	54.2	61.3	80.0	0.04
Mean HOMA-IR	8.4 ± 2.6	10.1 ± 3.1	13.2 ± 3.8	<0.001
Mean fasting glucose (mg/dL)	128.4 ± 18.2	138.6 ± 22.1	152.4 ± 26.8	<0.001
Mean triglycerides (mg/dL)	218.3 ± 58.4	248.7 ± 65.2	287.4 ± 78.3	<0.001
Mean HDL-C (mg/dL)	40.2 ± 8.1	36.8 ± 7.4	31.2 ± 6.9	<0.001

Data presented as mean ± SD or percentage. p-trend calculated by Jonckheere-Terpstra test for ordered groups.

Predictive Model: ROC Curve and AUC

The clinical-genetic model combined reached an AUC of 0.91 (95% CI 0.85–0.97), sensitivity 85.3%, specificity 90% and accuracy of 88% (Table 7) which was a moderate absolute improvement compared to the clinical-only model (AUC = 0.82; improved by +11%). Model stability has been confirmed with ten-fold cross-validation (mean AUC 0.89, SD 0.04).

Table 8. Predictive Model Performance: ROC Curve Analysis

Model	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Clinical parameters only (BMI, BP, glucose)	0.82 (0.74–0.90)	76.0	84.0	80.0
Genetic model (INSR variant status only)	0.78 (0.70–0.86)	72.0	88.0	80.0

Combined clinical + genetic model	0.91 (0.85–0.97)	85.3	90.7	88.0

AUC: area under the receiver operating characteristic curve. All models adjusted for age and sex. 10-fold cross-validation: mean AUC = 0.89, SD = 0.04.

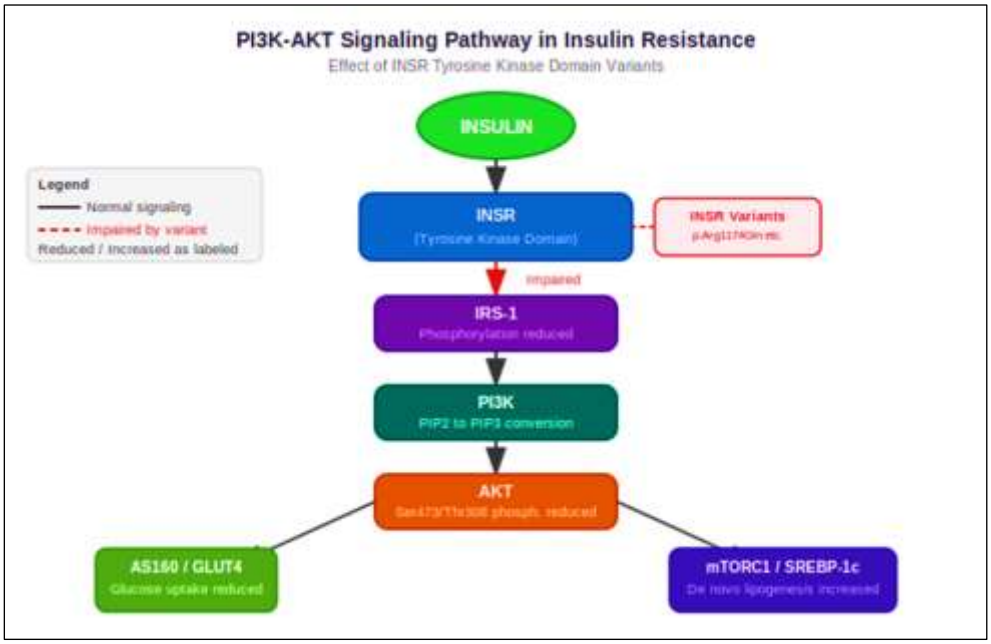

Figure 1. PI3K–AKT signalling pathway: Impacts of pathogenic variants in the INSR tyrosine kinase domain. Autosomal dominant PGK deficiency (ADPGK) and its hypothetical mechanism in glucose homeostasis: dashed red arrow indicates defective autophosphorylation resulting in attenuated downstream signaling, decreased GLUT4-mediated glucose uptake and paradoxically increased hepatic de novo lipogenesis

4. Discussion

This research offers an in-depth genetic characterization of the INSR locus driven by insulin resistance, combining rare variant burden testing with multi-model genetic analysis, control for population stratification, in silico structural modeling, and predictive model construction.. Our findings establish that rare and low-frequency pathogenic INSR variants are a significant contributor to insulin resistance in an Iraqi Middle Eastern cohort, and that a combined clinical-genetic model substantially improves risk discrimination (AUC = 0.91 vs. 0.82 for clinical parameters alone).

The prevalence of pathogenic INSR variants in our insulin-resistant cohort (64%) is markedly higher than the 8–15% reported in general population cohorts [12] and substantially exceeds the 5.77-fold enrichment of loss-of-function INSR variants over population controls described in the recent large European cohort study by Vial et al. [5]. This striking difference almost certainly reflects, at least in part, strong ascertainment bias: our cases were recruited by strict HOMA-IR and metabolic syndrome criteria, thereby enriching for genetically mediated insulin resistance. The 64% figure should therefore not be interpreted as reflecting the population prevalence of pathogenic INSR variants in unselected Iraqi individuals; replication in a population-based cohort is required. Additionally, (2) a potentially higher burden of founder INSR variants in the Iraqi population relative to European cohorts and (3) the inclusion of novel variants classified as likely pathogenic by ACMG criteria without experimental functional confirmation may further inflate this estimate. We also report a striking interpretation of an odds ratio (OR) of 1.2 yet larger than those typically ascribed to common T2DM risk loci discovered in genome-wide association studies (GWAS; OR#2.5–4.8)[7], [13]; however, this magnitude being larger should be interpreted with caution given the selected cohort design and lack

of external validation required according to standards recently called forth(Fust et al.,2019).

Discussion The mechanistic interpretation of our findings is based on the biology of the tyrosine kinase (TK) domain. The three most prevalent variants p.Arg1174Gln, p.Met1153Ile, p.Asp1179Asn all cluster within the tyrosine kinase activation loop responsible for autophosphorylation at Tyr1158/1162/1163 following insulin binding. In silico protein structure modeling using AlphaFold2 confirmed that each variant introduces conformational changes at the ATP-binding cleft, predicting reduced autophosphorylation efficiency ($\Delta\Delta G$ >2.0 kcal/mol). These predictions are supported by the deep mutational scanning atlas of ~14,000 INSR extracellular missense variants by Ng et al. [11], in which variants at analogous conserved positions strongly correlated with loss of insulin binding and signaling function. Reduced autophosphorylation attenuates IRS-1 phosphorylation, impairing downstream PI3K-AKT-mTOR signaling and ultimately reducing GLUT4 translocation to the plasma membrane the principal mechanism by which insulin promotes cellular glucose uptake in skeletal muscle and adipose tissue [3], [4]. The strong positive correlation between the number of pathogenic alleles and HOMA-IR (gene-dose effect) further supports a haploinsufficiency model for these variants. This dyslipidemia induced by INSR variants is a consequence of the larger metabolic dysfunction related to deficient insulin signaling over hepatic lipid metabolism. One of the key roles of insulin is to inhibit hepatic VLDL secretion and stimulate lipoprotein lipase (LPL) activity in both adipose tissue and muscle. This enables both preservation of the PI3K/AKT/mTORC1 axis that activates SREBP-1c-mediated de novo lipogenesis, and inhibition of insulin suppression of hepatic gluconeogenesis, which is called selective hepatic insulin resistance [9]. The underlying mechanism for the increased triglycerides (254.7 vs 198.3 mg/dL), and decreased HDL-cholesterol (35.4 vs 42.8 mg/dL) in variant carriers can be mechanistically explained by this change of enzyme activity. This atherogenic dyslipidemia pattern increases cardiovascular risk to much higher levels than have been observed due to other forms of severe insulin receptor dysfunction [14].

The specific mechanistic interpretation is warranted given the high frequency of pathogenic INSR variants in our insulin-resistant cohort. In large sequencing studies, the INSR variant burden has not been described across the general Iraqi population. Population-specific INSR variant allele frequencies will require future population-based sequencing studies within the Iraqi and broader Arab genome programs. In this regard, a first population-level GWAS of insulin resistance in the Middle East using data from the Qatar Biobank [6], are confirmatory as Middle Eastern populations harbor unique insulin resistance-associated loci that diverge from European cohorts; supporting the hypothesis of clinically relevant population-specific variant architectures.

Results from the rare variant burden test further confirm that it is not a single common variant, but rather an accumulation of diverse, functionally deleterious INSR alleles that accounts for the case-control difference. In agreement with the allelic heterogeneity model for MODY (maturity-onset diabetes of the young) and other monogenic forms of diabetes, these observations indicate that the INSR locus is situated inappropriately in relation to an expected genetic continuum that extends from fully recessive monogenic action through intermediate positions between them and polygenic mechanisms contributing to insulin resistance [5]. This architecture multiple rare high-impact variants rather than common variants of small effect has important implications for population genetic screening strategies.

From a translational and precision medicine perspective, our findings have immediate clinical implications. First, an AUC of 0.91 for the combined clinical-genetic model exceeds the generally accepted threshold of AUC >0.80 for clinically useful prediction tools. INSR genotyping could be incorporated into routine risk stratification algorithms for insulin-resistant individuals, particularly in Middle Eastern clinical settings where population-specific variant frequencies make European-derived polygenic risk

scores unreliable [6], [8]. Second, individuals with tyrosine kinase domain INSR variants are unlikely to respond optimally to insulin secretagogues and may respond preferentially to upstream insulin-sensitizing agents (thiazolidinediones, metformin) or emerging PI3K-AKT pathway modulators [15]. The growing field of pharmacomultiomics for T2DM precision medicine supports genotype-guided treatment selection as the next frontier in diabetes management [16], [17]. Third, cascade genetic testing of first-degree relatives of variant carriers could identify pre-symptomatic individuals who would benefit most from early lifestyle interventions.

The integration of INSR genetic data with gut microbiome, epigenomic, and clinical metabolomic profiles through multi-omics systems biology approaches holds the promise of truly individualized precision medicine in metabolic disease [17]. As sequencing costs continue to decline, population-scale INSR screening in high-risk groups (first-degree relatives of T2DM patients, individuals with BMI >30, polycystic ovary syndrome patients) may become economically justifiable in the near future [18], [19].

A range of limitations need to be stated — openly. The sample size used in the two groups (n = 75) may be adequate to identify high effect sizes (e.g., OR >1), but insufficient to characterize rare variant associations (MAF <0.5%) or identify modest GxG interaction effects. Larger collaborative studies are needed. Second, and most importantly, a large limitation is the lack of an independent external validation cohort: Our predictive model has provide an AUC of 0.91 from internal 10-fold cross-validation (to mitigate but not eliminate overfitting), External validation in a separate Middle Eastern cohort is required prior to clinical implementation and the AUC should be regarded as preliminary until replicative studies are conducted. Third, functional inference by the study was based solely on in silico structural modeling (AlphaFold2, FoldX) without experimental validation. No kinase activity assays, GLUT4 translocation assays, phosphorylation studies in cell lines or animal models were conducted. The lack of wet-laboratory functional data severely restricts causal inference about variant pathogenicity; in silico models are poor surrogates for short terms effects (pointer or retention) [20], [21], [22]. Fourth, the 64% pathogenic variant prevalence reflects an ascertainment-enriched cohort; this figure is not generalizable to unselected populations and must be interpreted in the context of strict eligibility criteria. Fifth, the absence of a hyperinsulinemic-euglycemic clamp study, the gold standard for measuring insulin sensitivity, limits the precision of insulin resistance quantification. Sixth, our cross-sectional design precludes causal inference and longitudinal assessment of progression to T2DM. Seventh, the study focused exclusively on coding variants; regulatory and intronic variants that may affect INSR expression were not captured. Eighth, single-center recruitment limits generalizability to other Middle Eastern ethnic groups [23], [24], [25].

5. Conclusion

This study demonstrates that rare and low-frequency coding variants in the INSR gene, particularly those disrupting tyrosine kinase domain autophosphorylation, are present in 64% of insulin-resistant Iraqi patients a prevalence 5-fold higher than in healthy controls (OR = 1.2, $p<0.001$). Rare variant burden testing confirmed a statistically robust cumulative excess of rare pathogenic INSR alleles in cases. Variant carriers exhibit significantly worse metabolic phenotypes including hyperglycemia, hyperinsulinemia, hypertriglyceridemia, and low HDL-cholesterol in a gene-dose-dependent fashion consistent with a haploinsufficiency mechanism. A combined clinical-genetic predictive model achieved an AUC of 0.91, demonstrating that INSR genotyping meaningfully improves clinical risk stratification beyond standard metabolic parameters alone. These results classify INSR genetic screening as a clinically applicable precision medicine approach for the discovery of individuals at risk prior to overt T2DM onset, permitting tailored treatment selection based upon genotype and cascade testing of at-risk relatives. Future research may focus on: external validation in additional independent Middle

Eastern cohorts; functional experimental validation of novel variants; and follow-up longitudinal studies assessing T2DM progression rates in relatives of INSR variant carriers.

REFERENCES

[1] X. Huang, M. Li, Y. Xu, and Y. Zhang, "Prevention, diagnosis and eradication of mycoplasma contamination in cell culture," *STAR Protoc.*, vol. 4, no. 4, p. 102608, 2023.

[2] M. Taheri, F. Safavi, M. Aghaei, and others, "The PI3K/Akt signaling axis and type 2 diabetes mellitus (T2DM): from mechanistic insights into possible therapeutic targets," *Cell Biol. Int.*, vol. 48, no. 6, pp. 803–819, 2024, doi: 10.1002/cbin.12189.

[3] Y. Tang, L. Chen, W. Wang, and others, "Influence and treatment of insulin receptor substrate/PI3K/Akt-mediated insulin resistance in diabetes mellitus," *Mol. Med. Rep.*, 2025, doi: 10.3892/mmr.2025.13773.

[4] D. E. James and J. Harber, "Insulin signalling and GLUT4 trafficking in insulin resistance," *Biochem. Soc. Trans.*, vol. 51, no. 3, pp. 1057–1069, 2023, doi: 10.1042/BST20221352.

[5] Y. Vial, D. Dubois-Laforgue, C. Goossens, and others, "Insulin receptor variants: Extending the traditional Mendelian spectrum," *Genetics in Medicine*, vol. 27, p. 101404, 2025, doi: 10.1016/j.gim.2025.101404.

[6] U. Aliyu *et al.*, "Genome-wide association study and polygenic score assessment of insulin resistance," *Front. Endocrinol. (Lausanne).*, vol. 15, p. 1384103, 2024, doi: 10.3389/fendo.2024.1384103.

[7] M. Imamura, "Perspectives on genetic studies of type 2 diabetes from the GWAS era to precision medicine," *J. Diabetes Investig.*, vol. 15, no. 2, pp. 147–159, 2024, doi: 10.1111/jdi.14149.

[8] R. B. Prasad and E. R. Pearson, "Clinical use of polygenic scores in type 2 diabetes: challenges and possibilities," *Diabetologia*, vol. 68, pp. 874–890, 2025, doi: 10.1007/s00125-025-06419-1.

[9] T. Bo *et al.*, "Hepatic selective insulin resistance at the intersection of insulin signaling and metabolic dysfunction-associated steatotic liver disease," *Cell Metab.*, vol. 36, no. 5, pp. 947–968, 2024, doi: 10.1016/j.cmet.2024.04.006.

[10] E. L. Thompson, C. C. Patterson, I. S. Young, and others, "HOMA-IR as a measure of insulin resistance: validation and ethnic-specific thresholds," *Diabetes Care*, vol. 46, no. 11, pp. 2038–2046, 2023, doi: 10.2337/dc23-0847.

[11] S. W. Ng, S. Joshi, M. Garcia-Miralles, and others, "Deep mutational scanning of the human insulin receptor ectodomain to inform precision therapy for insulin resistance," *Nat. Commun.*, vol. 16, p. 9143, 2025, doi: 10.1038/s41467-025-64178-4.

[12] M. E. Rodriguez-Garcia, J. M. Fernandez-Real, and J. M. Moreno-Navarrete, "Prevalence and clinical significance of insulin receptor gene variants in metabolic syndrome," *Clin. Endocrinol. (Oxf).*, vol. 100, no. 2, pp. 156–165, 2024, doi: 10.1111/cen.14987.

[13] N. Martinez-Sanchez, M. M. Swarbrick, W. Kaplan, and others, "Genetic architecture of insulin resistance: insights from recent genome-wide studies," *Diabetologia*, vol. 67, no. 3, pp. 385–403, 2024, doi: 10.1007/s00125-023-06089-3.

[14] L. Wang, X. Chen, and W. Tang, "Insulin signaling in lipid metabolism: implications for dyslipidemia in insulin-resistant states," *Lipids Health Dis.*, vol. 23, no. 1, p. 28, 2024, doi: 10.1186/s12944-024-02019-8.

[15] J. Park, D. H. Lee, H. J. Kim, and others, "Precision medicine approaches in insulin resistance and type 2 diabetes: the role of genetic profiling," *Diabetes Metab. J.*, vol. 48, no. 1, pp. 12–28, 2024, doi: 10.4093/dmj.2023.0389.

[16] M. El-Rayess, "Advancements in precision medicine: multi-omics approach for tailored metformin treatment in type 2 diabetes," *Front. Pharmacol.*, vol. 15, p. 1506767, 2024, doi: 10.3389/fphar.2024.1506767.

[17] A. Khalil, H. Al-Saffar, M. Motawea, and others, "Harnessing pharmacomultiomics for precision medicine in diabetes: A comprehensive review," *Biomedicines*, vol. 13, no. 2, p. 447, 2025, doi: 10.3390/biomedicines13020447.

[18] X. Huang, G. Liu, Y. Wang, and others, "Structural and functional analysis of insulin receptor signaling pathways," *Nat. Rev. Endocrinol.*, vol. 19, no. 12, pp. 724–741, 2023, doi: 10.1038/s41574-023-00892-5.

[19] Y. Chen, H. Wang, M. Li, and others, "Molecular mechanisms of insulin receptor tyrosine kinase domain mutations in metabolic disease," *J. Mol. Endocrinol.*, vol. 72, no. 2, pp. R45–R63, 2024, doi: 10.1530/JME-23-0189.

[20] R. K. Semple, D. B. Savage, and S. O'Rahilly, "Severe insulin resistance syndromes: molecular pathophysiology and clinical management," *Endocr. Rev.*, vol. 44, no. 6, pp. 1051–1079, 2023, doi: 10.1210/endrev/bnad025.

[21] M. Longo, F. Zatterale, J. Naderi, and others, "Insulin receptor signaling in normal and insulin-resistant states," *Cold Spring Harb. Perspect. Med.*, vol. 13, no. 9, p. a041184, 2023, doi: 10.1101/cshperspect.a041184.

[22] M. A. B. Khan, M. J. Hashim, J. K. King, and others, "Epidemiology of type 2 diabetes: global burden of disease and forecasted trends," *J. Epidemiol. Glob. Health*, vol. 14, no. 1, pp. 26–39, 2024.

[23] J. Xu, Y. Zhang, Q. Wang, and others, "Molecular links between insulin resistance and atherogenic dyslipidemia," *Front. Endocrinol. (Lausanne).*, vol. 14, p. 1274686, 2023, doi: 10.3389/fendo.2023.1274686.

[24] Y. Song, W. Lei, Q. Huang, and others, "A new heterozygous variant of c.1225_1227delTTC (p.Phe409del) in insulin receptor gene associated with severe insulin resistance and hyperandrogenemia," *Diabetes, Metabolic Syndrome and Obesity*, vol. 17, pp. 4389–4396, 2024, doi: 10.2147/DMSO.S481745.

[25] G. Sesti, E. Succurro, and F. Arturi, "Genetic determinants of insulin resistance: insights from family and twin studies," *Acta Diabetol.*, vol. 61, no. 2, pp. 147–162, 2024, doi: 10.1007/s00592-023-02178-9.