International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Role of INSR and MTHFR Gene Variants in PCOS: A Multi-Omics Review

Ilakiya T [1]*, Keerthana S L[1], Justin Packia Jacob[1]

[1]Department of Biotechnology, St. Joseph's College of Engineering, Chennai, Tamil Nadu 600119, India

*Corresponding author: ilakiya1909@gmail.com

Abstract - Polycystic Ovary Syndrome (PCOS) affects 8–13% of reproductive-age women globally, with prevalence reaching 22.5% in urban Indian cohorts. Two of the most consistently implicated genetic variants INSR rs1799817 and MTHFR C677T have almost always been studied in isolation. This review argues that studying them separately is the wrong approach. INSR rs1799817 impairs insulin receptor autophosphorylation, while MTHFR C677T reduces enzymatic activity by 35–70%, depleting S-adenosylmethionine (SAM) and elevating plasma homocysteine. These pathways intersect and amplify each other through what we term the insulin-folate axis. This review synthesises evidence published between 2019 and 2025 across five molecular layers genomics, transcriptomics, epigenomics, metabolomics, and proteomics to construct a unified mechanistic model. We evaluate multi-omics integration frameworks including MOFA+, Similarity Network Fusion, iCluster+, and OmicsAnalyst, identify South Asian populations as critically underrepresented in this research, and propose a cost-effective tiered diagnostic workflow based on Tetra-ARMS PCR genotyping.

Key Words: *PCOS, INSR rs1799817, MTHFR C677T, insulin-folate axis, multi-omicsintegration, insulin resistance.*

1. INTRODUCTION

1.1 Clinical and Epidemiological Burden of PCOS

Polycystic Ovary Syndrome is the most prevalent endocrine disorder of reproductive-age women globally, with prevalence estimates ranging from 8 to 13% depending on the diagnostic criteria applied and the population studied [1]. Population-specific surveys in urban India have reported figures as high as 22.5%, a disparity that cannot be explained by differences in diagnostic ascertainment alone and likely reflects genuine population-level differences in genetic risk architecture and environmental exposure [1]. Clinically, PCOS is defined by the Rotterdam 2003 consensus criteria, which require at least two of three cardinal features: hyperandrogenism, oligo-anovulation, and polycystic ovarian morphology on ultrasound [2]. Insulin resistance affects 50–70% of PCOS patients irrespective of body weight, whilst dyslipidaemia, non-alcoholic fatty liver disease, type 2 diabetes, and cardiovascular disease represent downstream consequences that substantially reduce healthy life expectancy [2].

In lower-middle-income settings such as India, NGS-based molecular diagnostics remain largely inaccessible outside tertiary academic centres. The diagnostic delay for PCOS routinely extends to several years, with considerable downstream cost to patients and health systems [1,2]. This provides the central motivation for the present review: the need for a mechanistically grounded, computationally accessible, and economically deployable molecular framework that can function within South Asian clinical constraints.

1.2 Genetic Architecture of PCOS and the Problem with Single-Gene Studies

Genome-wide association studies (GWAS) across European, Chinese, and South Asian cohorts have catalogued more than twenty reproducible susceptibility loci, encompassing genes involved in gonadotropin signalling (LHCGR, FSHR), cell cycle regulation (THADA), vesicle trafficking (DENND1A), and transcriptional control (YAP1), amongst others [3]. Yet the aggregate explained heritability across these loci remains modest. This reflects the fundamental limitations of single-variant additive models when applied to a polygenic, clinically heterogeneous syndrome [3].

Two candidate genes have accumulated convergent and reproducible evidence across multiple continents and ethnic groups. The Insulin Receptor gene (INSR), located at chromosome 19p13.2, encodes the heterotetramer whose intracellular beta-subunit tyrosine kinase domain transduces the insulin signal through IRS-1-mediated phosphorylation cascades. The synonymous rs1799817 variant — a C-to-T transition at codon 1058 in exon 17 — has been associated with impaired receptor autophosphorylation across Indian,

Turkish, and Iraqi PCOS cohorts [4,5]. Methylenetetrahydrofolate Reductase (MTHFR), located at chromosome 1p36.22, encodes the rate-limiting enzyme of the folate cycle. The C677T variant reduces enzymatic activity by 35% in CT heterozygotes and 60–70% in TT homozygotes, elevating plasma homocysteine and depleting SAM availability for methylation reactions [6,7].

1.3 The Insulin-Folate Axis

The central argument of this review is that the pathways governed by INSR and MTHFR do not operate independently. Their biochemical intersection — which we term the insulin-folate axis — is bidirectional and self-amplifying. Impaired INSR-mediated glucose transport starves the folate cycle of carbon substrate. Homocysteine accumulating from MTHFR dysfunction then oxidatively damages the INSR tyrosine kinase domain, compounding the signalling deficit [7]. Each variant makes the other worse. No study has examined both in a single matched cohort. This review identifies that gap, characterises its mechanistic basis across five molecular layers, and proposes how it can be closed.

1.4 Why Multi-Omics Integration Is Necessary

A bidirectional, cross-pathway interaction like the insulin-folate axis cannot be captured by single-platform analysis. Between 2019 and 2025, multi-omics integration methodology has matured substantially. Frameworks including MOFA+, Similarity Network Fusion (SNF), iCluster+, and OmicsAnalyst can now decompose shared biological variation across genomic, transcriptomic, epigenomic, metabolomic, and proteomic data from the same patients simultaneously [8,9,10,11]. Despite this methodological maturity, no existing review frames INSR and MTHFR as co-drivers of PCOS through convergent, omics-validated mechanisms. This review addresses that gap.

2. GENOMIC EVIDENCE: INSR AND MTHFR AS SUSCEPTIBILITY LOCI

2.1 Population Genetics and the Limits of GWAS

INSR and MTHFR have not consistently attained genome-wide significance thresholds ($p < 5\times10^{-8}$) in the largest multi-ancestry GWAS analyses. This is not evidence of biological irrelevance. Three factors explain the discordance: small-to-moderate effect sizes; substantial allele frequency differences between European and South Asian cohorts; and phenotypic heterogeneity that dilutes statistical signal when clinically distinct subtypes are analysed together [3,4].

A landmark study in lean Indian women reported that the C/T polymorphism at His1058 (rs1799817) in exon 17 of INSR was significantly associated with PCOS ($\chi^2 = 8.493$, $p = 0.004$), correlating specifically with elevated fasting insulin ($p = 0.02$), HOMA-IR ($p = 0.005$), and free androgen index ($p = 0.03$) [4]. For MTHFR, a meta-analysis of 22 independent studies encompassing 2,405 PCOS cases and 2,419 controls confirmed that the T allele of C677T is associated with increased PCOS risk, with an odds ratio of 1.40 (95% CI: 1.27–1.53) and effect sizes notably larger in Asian versus Caucasian populations [6].

2.2 Functional Consequences of INSR rs1799817

Although rs1799817 produces a synonymous His1058 substitution at the amino acid level, it carries measurable functional weight. Structural modelling and kinase assay data indicate that the T-allele reduces receptor autophosphorylation efficiency, impairing activation of IRS-1 at Tyr632, PI3K, and AKT-mediated GLUT4 translocation [4,5]. A 2022 cohort study found the C allele frequency at 0.46 in PCOS patients versus 0.30 in controls, with specific genotype combinations associated with significantly elevated BMI, triglycerides, and fasting glucose [38]. In ovarian theca cells, impaired receptor signalling fails to suppress CYP17A1 expression, directly linking INSR genotype to the hyperandrogenic phenotype [4,14].

2.3 Functional Consequences of MTHFR C677T

The Ala222Val substitution produced by C677T destabilises FAD cofactor binding and reduces catalytic efficiency in a dose-dependent manner: approximately 35% in CT heterozygotes and 60–70% in TT homozygotes [6,7]. The resulting homocysteine accumulation has been mechanistically linked to PCOS through three converging pathways: endothelial dysfunction and microvascular insulin resistance; NF-κB-mediated pro-inflammatory signalling; and direct inhibition of DNA methyltransferase activity through elevated S-adenosylhomocysteine (SAH) [7,13]. A 2025 case-control study of Georgian women confirmed significant associations between MTHFR C677T and adverse pregnancy outcomes, with TT homozygotes demonstrating the most severe metabolic and reproductive phenotype [13].

3. TRANSCRIPTOMIC EVIDENCE

3.1 INSR as a Hub Gene in the Steroidogenic Transcriptional Module

RNA-seq studies of granulosa cells from PCOS patients have consistently identified significant downregulation of INSR transcript levels. Weighted Gene Co-expression Network Analysis (WGCNA) reveals that INSR emerges as a hub gene within a transcriptional module that co-expresses with the steroidogenic enzymes CYP11A1, HSD3B2, and STAR in PCOS granulosa cells [23]. This means that INSR expression changes propagate into androgen biosynthesis through transcriptional network effects — not simply through linear signal transduction.

3.2 MTHFR: Cell-Type Specificity and One-Carbon Network Suppression

Single-cell RNA sequencing has revealed that MTHFR transcriptional activity is lowest specifically in cumulus-oocyte complexes, particularly in patients carrying the TT genotype [14]. This matters because the cumulus-oocyte complex depends critically on MTHFR-supplied methyl groups during oocyte maturation. Differential expression analysis has further identified coordinate downregulation of MTHFR alongside SHMT1, MTR (methionine synthase), and DNMT3A in PCOS granulosa cells, consistent with suppression of the entire one-carbon metabolic network [24,34]. Notably, this transcriptional reprogramming has been observed in some datasets independently of genotype, suggesting that epigenetic mechanisms can propagate one-carbon suppression even in wild-type individuals.

4. EPIGENOMIC ARCHITECTURE: THE MTHFR-SAM-DNMT AXIS

Epigenomics provides the most direct mechanistic link between MTHFR genotype and PCOS phenotype. MTHFR catalyses the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which supplies the methyl group for remethylation of homocysteine to methionine. Methionine is adenosylated to SAM, the universal methyl donor for all DNA methyltransferases (DNMTs) [7]. Reduced MTHFR activity in C677T carriers simultaneously depletes SAM and elevates SAH, impairing all three classes of methyltransferases.

4.1 Global Hypomethylation, Focal Hypermethylation, and Androgen Gene De-Repression

SAM depletion driven by MTHFR C677T produces a layered epigenomic landscape in PCOS. Global CpG hypomethylation liberates chromatin at repeat elements and transcription factor binding sites, activating latent inflammatory programmes and disrupting genome stability [15,16]. Most consequentially, promoter hypomethylation at CYP11A1 and the androgen receptor (AR) gene — confirmed in TT genotype carriers — provides a direct epigenetic mechanism by which a folate cycle variant increases steroidogenic enzyme expression and androgen receptor sensitivity [7,15]. Restoration of methylation capacity through 5-methyltetrahydrofolate supplementation would be predicted to reverse this promoter hypomethylation and reduce androgen excess. This provides a molecularly targeted rationale for folate supplementation that is substantially more specific than conventional nutritional justification.

4.2 Transgenerational Transmission and Causal Evidence

A 2022 study demonstrated global DNA hypomethylation in both PCOS-affected women and their daughters, and confirmed that SAM supplementation substantially normalised PCOS-like traits in a transgenerational mouse model [16]. A 2025 Mendelian randomisation study provided causal epidemiological evidence that disrupted folate metabolism increases PCOS risk, whilst simultaneously demonstrating that PCOS may causally reduce vitamin B12 levels [17]. This bidirectional causal relationship substantially strengthens the biological plausibility of the MTHFR-SAM-DNMT axis as a therapeutic target.

5. METABOLOMIC PROFILING: THE INSULIN-FOLATE AXIS AS A BIOCHEMICAL SIGNATURE

Metabolomics provides the most functionally direct readout of the combined biochemical impact of INSR and MTHFR variants. LC-MS/MS and NMR spectroscopy have been applied extensively to PCOS metabolomics, with converging findings across diverse ethnic populations [18,19]. Pathway enrichment analyses using MetaboAnalyst 5.0 [20] have consistently highlighted three disrupted superpathways in PCOS plasma: one-carbon metabolism, glycolysis and gluconeogenesis, and steroid hormone biosynthesis — precisely the three superpathways predicted to be

disrupted by co-occurrence of INSR rs1799817 and MTHFR C677T.

5.1 The Homocysteine-Insulin Nexus

The dominant metabolomic signature in PCOS encompasses elevated plasma homocysteine, reduced circulating folate and vitamin B12, hyperinsulinaemia, elevated HOMA-IR, dysregulated branched-chain amino acids, and altered phospholipid profiles consistent with oxidative membrane damage [18,19,21]. Elevated homocysteine modifies cysteine residues within the INSR tyrosine kinase domain through thiol oxidation, reducing receptor autophosphorylation efficiency independently of the rs1799817 genotype, thereby compounding the intrinsic signalling deficit the variant already imposes [7,19]. This creates a self-amplifying cycle: MTHFR insufficiency raises homocysteine → homocysteine impairs INSR function → worsening insulin resistance → compensatory hyperinsulinaemia → CYP17A1 overexpression in theca cells → androgen excess → further metabolic dysregulation.

5.2 The Most Significant Unresolved Gap in PCOS Metabolomics

To date, no published study has examined the combined metabolomic consequence of rs1799817 and C677T co-occurrence in a single well-characterised, genotyped cohort. The additive or synergistic impact of dual variant carriage on plasma homocysteine, HOMA-IR, and one-carbon pathway metabolite ratios remains formally unquantified. This is the most significant empirical gap at the metabolomic layer of the insulin-folate axis model. Homocysteine immunoassay and HOMA-IR together provide the most accessible proxy measurements for MTHFR and INSR functional status, respectively, in resource-limited settings [20,21].

6. PROTEOMIC EVIDENCE

6.1 The INSR Proteomic Signature

Proteomic analyses have confirmed downregulation of total INSR protein abundance and reduced phosphorylation of IRS-1 at Tyr632, AKT at Ser473, and impaired GLUT4 membrane translocation in rs1799817 T-allele carriers [22,23]. A 2022 SWATH-MS study identified differential abundance of proteins within the mTOR, FOXO1, and AMPK signalling hubs [22], confirming that INSR variant-driven deficits propagate through the metabolic regulatory network rather than remaining confined to the immediate insulin-signalling cascade.

6.2 The MTHFR Proteomic Signature and Convergence Nodes

Proteomic analyses have identified reduced abundance of methionine synthase (MTR), methylenetetrahydrofolate dehydrogenase (MTHFD1), and serine hydroxymethyltransferase (SHMT1) in PCOS samples from C677T carriers, consistent with coordinate suppression of the entire one-carbon enzyme complex [24]. Protein-protein interaction network analysis using STRING v12.0 [25] has confirmed that INSR and MTHFR-encoded proteins participate in overlapping regulatory hubs involving AMPK, mTOR, DNMT3A, and FOXO1. AMPK and mTOR are the nodal proteins at which the two pathways physically converge — a finding with direct therapeutic implications as both are established pharmacological targets.

7. BIOINFORMATICS AND MULTI-OMICS INTEGRATION

7.1 The Analytical Challenge

Each molecular layer is characterised by fundamentally different data structures: genomic data are discrete; transcriptomic data are count-distributed with overdispersion; metabolomic data are continuous with log-normal distributions; DNA methylation beta-values are bounded between zero and one. Analysing these layers independently and then narratively combining the results risks missing cross-layer relationships detectable only through simultaneous integration [8,9].

7.2 Integration Frameworks

MOFA+ (Multi-Omics Factor Analysis v2) is the most widely adopted framework for unsupervised multi-omics integration [10]. Its probabilistic latent factor model decomposes shared and layer-specific variation across matched datasets simultaneously, identifying latent biological processes without requiring prior biological knowledge. Applied to a matched PCOS cohort, MOFA+ would be expected to identify a latent factor loading on INSR expression, MTHFR methylation status, plasma homocysteine, and downstream androgen biosynthesis gene expression — an unbiased, data-driven validation of the insulin-folate axis.

Similarity Network Fusion (SNF) constructs sample-similarity networks for each omics layer and fuses them into a consensus network for patient stratification [11]. Applied to PCOS, SNF could identify two molecularly coherent subtypes: an INSR-dominant metabolic subtype and an MTHFR-dominant epigenomic subtype. These subtypes carry distinct therapeutic predictions —

the first may respond preferentially to metformin or GLP-1 receptor agonists, the second represents the primary candidate for targeted 5-methyltetrahydrofolate supplementation [17].

7.3 Layer-Specific Computational Pipelines

For transcriptomics, DESeq2 [42] and edgeR [43] remain the standards for differential expression analysis, with STAR or HISAT2 for read alignment to GRCh38. For epigenomics, the Bismark aligner [44] handles bisulfite-treated read alignment; minfi and ChAMP [45] provide differentially methylated region identification from Illumina EPIC 850K array data. MetaboAnalyst 5.0 [20] remains the most comprehensive publicly accessible platform for metabolomics pathway enrichment and biomarker discovery. For proteomics, MaxQuant and Perseus constitute the standard pipeline for label-free quantification. For the genotyping workflow central to this review, Tetra-ARMS PCR enables simultaneous discrimination of wild-type, heterozygous, and homozygous mutant genotypes for both INSR rs1799817 and MTHFR C677T in a single reaction [12], with results available within 4–6 hours.

Table-1: Multi-omics evidence for INSR rs1799817 and MTHFR C677T in PCOS pathogenesis (studies published 2019–2025).

Omics Layer	Key Genes / Markers	Platform	Principal Findings
Genomics	INSR rs1799817, MTHFR C677T	GWAS, Tetra-ARMS PCR, PCR-RFLP	INSR T-allele associated with PCOS in lean Indian women (χ^2=8.493, p=0.004); MTHFR C677T OR 1.40 (95% CI 1.27–1.53); effects larger in Asian cohorts
Transcriptomics	INSR mRNA, CYP11A1, AR, DNMT3A, SHMT1	RNA-seq, scRNA-seq, DESeq2, WGCNA	INSR downregulated as hub gene in steroidogenic module; MTHFR lowest in cumulus-oocyte complexes of TT carriers; one-carbon enzyme network co-suppressed
Epigenomics	CpG methylation at CYP11A1, AR, INSR loci	RRBS, Illumina EPIC 850K, Bismark, minfi	MTHFR C677T drives global SAM depletion and CpG hypomethylation; CYP11A1/AR promoter hypomethylation amplifies androgen synthesis; transgenerational transmission confirmed in mouse model
Metabolomics	Homocysteine, HOMA-IR, BCAAs, folate, insulin	LC-MS/MS, NMR, MetaboAnalyst 5.0	Elevated Hcy in MTHFR TT carriers oxidatively impairs INSR; three disrupted superpathways align precisely with dual-variant predictions; dual-variant metabolomic cohort study absent from literature
Proteomics	INSR, IRS-1 pTyr632, AKT pSer473, MTR, SHMT1	iTRAQ, SWATH-MS, STRING v12.0	Reduced INSR abundance and impaired downstream phosphorylation; one-carbon enzyme complex co-downregulated; AMPK and mTOR identified as protein-level convergence

			nodes
Multi-omics	INSR–MTHFR axis	MOFA+, SNF, iCluster+, OmicsAnalyst	Convergent dysregulation across all five layers defines the insulin-folate axis as the central shared pathogenic hub; patient subtype stratification predicts differential therapeutic responses

Table -2: Computational tools for multi-omics integration in PCOS research.

Tool	Omics Layer	Function	Application in This Review
MOFA+	Multi-omics	Probabilistic latent factor decomposition	Identify shared variation between genomic, epigenomic, and metabolomic layers; data-driven validation of the insulin-folate axis
SNF	Multi-omics	Patient-similarity network fusion	Stratify PCOS into INSR-dominant and MTHFR-dominant molecular subtypes for genotype-informed therapy
iCluster+	Genomics + Transcriptomics	Joint latent variable clustering	Discover molecularly coherent patient subtypes for precision stratification
DESeq2 / edgeR	Transcriptomics	Differential expression analysis	Compare INSR and MTHFR expression in PCOS vs. control granulosa and theca cells
Bismark / minfi / ChAMP	Epigenomics	Bisulfite alignment and DMR analysis	Identify CYP11A1 and AR promoter DMRs linked to MTHFR C677T in PCOS tissue
MetaboAnalyst 5.0	Metabolomics	Pathway enrichment, biomarker discovery	Map one-carbon and insulin-signalling disruption; characterise homocysteine-HOMA-IR signatures
STRING / Cytoscape	Proteomics	Protein interaction network analysis	Map INSR–MTHFR interaction hubs; confirm AMPK and mTOR as convergence nodes
Tetra-ARMS PCR + in silico RFLP	Genomics	Allele-specific PCR genotyping + computational validation	Rapid, cost-effective genotyping of both variants without NGS; deployable in South Asian clinical settings

8. RESEARCH GAPS AND THE SOUTH ASIAN EVIDENCE DEFICIT

8.1 The South Asian Population is Critically Underrepresented

The overwhelming majority of GWAS, transcriptomic, and metabolomic PCOS studies have been conducted in European or East Asian cohorts. South Asian populations, despite carrying the highest global PCOS burden, constitute a minority of multi-omics study cohorts [1,3]. INSR and MTHFR allele frequencies differ substantially between South Asian and Western populations, affecting the population attributable risk for each variant in Indian cohorts. The South Asian PCOS phenotype frequently presents as lean insulin resistance — the context in which INSR rs1799817 effects are strongest [4]. Environmental modifiers of one-carbon metabolism, including dietary folate and vitamin B12 availability in predominantly vegetarian diets, differ systematically between populations, modifying MTHFR C677T penetrance in ways that Western-derived estimates cannot capture.

8.2 The Missing Dual-Variant Cohort Study

The specific question of whether INSR rs1799817 and MTHFR C677T interact epistatically, act through convergent independent pathways, or compound each other's effects through the insulin-folate axis biochemistry has not been formally tested in any adequately powered single study. The absence of such data is the most significant unresolved empirical gap in the field.

8.3 A Tiered Diagnostic Approach for Resource-Limited Settings

A practical framework begins with Tetra-ARMS PCR genotyping for INSR rs1799817 and MTHFR C677T, complemented by in silico RFLP validation, HOMA-IR, and plasma homocysteine measurement for all patients. High-risk individuals identified by this first tier proceed to targeted transcriptomic profiling and selective Illumina EPIC 850K methylation analysis. Research-grade cohorts proceed to full five-layer multi-omics profiling with MOFA+ and SNF integration. This hierarchical workflow preserves the biological depth of multi-omics analysis while distributing cost across risk strata.

9. FUTURE PERSPECTIVES: TOWARDS INTEGRATIVE PRECISION MEDICINE

9.1 Machine Learning for Non-Linear Cross-Layer Interactions

Graph neural networks and attention-based transformer models trained on matched multi-omics matrices can capture non-linear cross-layer interactions that conventional linear integration methods cannot represent [8,9]. Applied to PCOS data integrating INSR genotype, MTHFR methylation status, plasma homocysteine, and downstream transcriptomic profiles, such models could generate molecular risk scores for disease severity and treatment response that substantively advance beyond symptom-based stratification.

9.2 Single-Cell Multi-Omics and the Ovarian Microenvironment

Platforms enabling simultaneous chromatin accessibility and gene expression measurement within individual cells — such as 10x Genomics Multiome ATAC + Gene Expression — would permit cell-type-specific mapping of INSR and MTHFR variant consequences within the heterogeneous ovarian microenvironment [40]. Combining ATAC-seq chromatin accessibility profiles with MTHFR methylation data would directly address whether C677T-driven SAM depletion physically opens chromatin at steroidogenic gene regulatory elements.

9.3 Genotype-Informed Therapeutic Stratification

The insulin-folate axis model generates specific and testable therapeutic predictions. Patients with INSR-dominant PCOS are predicted to respond preferentially to insulin sensitisers such as metformin and GLP-1 receptor agonists. Patients with MTHFR-dominant PCOS — TT homozygotes with marked hyperhomocysteinaemia, reduced plasma folate, and CYP11A1 promoter hypomethylation — are the primary candidates for targeted 5-methyltetrahydrofolate supplementation [17]. Dual-variant carriers may require combination strategies targeting both axes simultaneously.

9.4 The Path Forward for South Asian Research

For the South Asian clinical context, the path forward requires three coordinated steps: multi-institutional cohort formation with standardised recruitment protocols; biobanking infrastructure that collects matched blood, granulosa cells, and follicular fluid from the same patients; and capacity-building in computational biology within Indian research

institutions. The analytical frameworks reviewed here — MOFA+, SNF, MetaboAnalyst 5.0, OmicsAnalyst, and the R/Bioconductor ecosystem — are all freely accessible. The bottleneck is not methodology. It is matched, longitudinal, multi-omics PCOS data from South Asian cohorts with carefully characterised INSR and MTHFR genotypes.

10. LIMITATIONS

Several limitations of the present review must be acknowledged. First, this is a narrative review without a formally registered PRISMA search protocol; the synthesis reflects the authors' structured assessment of the literature rather than a reproducible systematic search. Second, the majority of studies cited are cross-sectional in design, precluding causal inference about temporal relationships between genotype, intermediate omics phenotypes, and clinical outcomes. Third, the functional evidence for INSR rs1799817 derives primarily from structural modelling and indirect kinase assay data; direct causal demonstration in isogenic human cell lines would substantially strengthen the mechanistic claims. Fourth, the insulin-folate axis is currently a theoretically grounded synthesis of evidence from independent studies. Its validity as a unified mechanistic framework awaits empirical testing in a single matched multi-omics cohort.

11. CONCLUSION

This review has constructed a five-layer multi-omics framework for understanding the converging pathogenic roles of INSR rs1799817 and MTHFR C677T in Polycystic Ovary Syndrome. At the genomic layer, both variants confer measurable susceptibility with effects amplified in Asian populations and the lean PCOS phenotype. At the transcriptomic layer, INSR functions as a hub gene within the steroidogenic co-expression module, whilst MTHFR suppression is concentrated in the cumulus-oocyte complex. At the epigenomic layer, MTHFR C677T depletes SAM, induces global DNA hypomethylation, and de-represses the CYP11A1 and androgen receptor promoters a direct epigenetic driver of the hyperandrogenic phenotype that may be transmitted transgenerationally. At the metabolomic layer, plasma homocysteine elevation and insulin resistance converge through the homocysteine-insulin nexus. At the proteomic layer, AMPK and mTOR are the protein-level convergence nodes.

INSR and MTHFR are not independent susceptibility factors that happen to co-occur in PCOS. They are mechanistically co-dependent nodes in a shared pathogenic network. Understanding PCOS through the lens of the insulin-folate axis, interrogated simultaneously across molecular layers, is not an analytical refinement. It is what the biology of this syndrome has always required.

DECLARATIONS

Funding: No external funding was received in the preparation of this manuscript.

Conflict of Interest: The authors declare no competing interests.

Ethical Approval: Not applicable (review article).

Data Availability: Not applicable.

Author Contributions: Ilakiya T and Keerthana SL conducted the literature review, synthesised the evidence, and drafted the manuscript. Dr. Justin Packia Jacob supervised the project, provided conceptual guidance, and critically reviewed the manuscript.

REFERENCES

1. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod. 2016;31(12):2841–2855.

2. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum Reprod. 2018;33(9):1602–1618.

3. Dapas M, Dunaif A. Deconstructing a syndrome: genomic insights into PCOS causal mechanisms and classification. Endocr Rev. 2022;43(6):927–965.

4. Saxena R, Jain M, Thomas S, Agrawal S, Singh K, Trivedi SS. Genetic variation in exon 17 of INSR is associated with insulin resistance and hyperandrogenemia among lean Indian women with polycystic ovary syndrome. J Genet. 2008;89(2):199–204.

5. Ozgul-Kilic SS, Kilic H, Gormus U. Investigation of associations between polycystic ovary syndrome and INSR gene polymorphisms rs2059806 and rs2252673. World J Diabetes. 2025;16(3):100121.

6. Li Y, Zhu H, Liu M, Zeng Z, Zeng Y, Xu X, Ye M. Significant association between methylenetetrahydrofolate reductase gene C677T polymorphism with polycystic ovary syndrome risk:

a meta-analysis update. Medicine (Baltimore). 2020;99(4):e18720.

7. Zhao J, Li X, Chen Q. Effects of MTHFR C677T polymorphism on homocysteine, insulin resistance, and oxidative stress in polycystic ovary syndrome: a systematic review. Reprod Biol Endocrinol. 2021;19(1):5.

8. Santos-Rodriguez G, Armenteros-Monterroso E, Cossio-Tejido I, et al. A technical review of multi-omics data integration methods: from classical statistical to deep generative approaches. Brief Bioinform. 2025;26(4):bbaf355.

9. Gomez-Cabrero D, Abugessaisa I, Maier D, et al. Multiomics research: principles and challenges in integrated analysis. BioDesign Res. 2024;6:0059.

10. Argelaguet R, Arnol D, Bredikhin D, Deloro Y, Velten B, Marioni JC, Stegle O. MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. Genome Biol. 2020;21(1):111.

11. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, Haibe-Kains B, Goldenberg A. Similarity network fusion for aggregating data types on a genomic scale. Nat Methods. 2014;11(3):333–337.

12. Al-Allawi N, Jubrael J, Al-Zihiry K. Detection of MTHFR C677T genotyping for spina bifida defects using ARMS-PCR technique. Ceylon Med J. 2025;66(3):9491.

13. Asanidze E, Kristesashvili J, Parunashvili N, et al. The relationship between homocysteine levels, MTHFR C677T and A1298C polymorphism, and pregnancy outcomes in Georgian women with polycystic ovary syndrome: a case-control study. Int J Reprod Biomed. 2025;23(1):25–34.

14. Zhao H, Zhang J, Cheng X, Nie X, He B. Insulin resistance in polycystic ovary syndrome across various tissues: an updated review of pathogenesis, evaluation, and treatment. J Ovarian Res. 2023;16(1):9.

15. Naseem A, Zahid M, Arshad K, et al. Differential methylation signatures associated with PCOS: a systematic review and in-silico analysis. Reprod Sci. 2025 [Epub ahead of print]. doi:10.1007/s43032-025-01976-7.

16. Abbott DH, Dumesic DA, Levine JE. Passing on PCOS: new insights into its epigenetic transmission. Cell Metab. 2022;35(2):196–197.

17. Chen M, Wang X, Liu Y, et al. One-carbon metabolism and risk of PCOS: insights from Mendelian randomization. J Ovarian Res. 2025;18(1):45.

18. Wörheide MA, Krumsiek J, Kastenmüller G, Arnold M. Multi-omics integration in biomedical research — a metabolomics-centric review. Anal Chim Acta. 2021;1141:144–162.

19. Murri M, Luque-Ramírez M, Insenser M, Ojeda-Ojeda M, Escobar-Morreale HF. Circulating markers of oxidative stress and polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod Update. 2013;19(3):268–288.

20. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021;49(W1):W388–W396.

21. Thornburgh S, Gaskins AJ. B vitamins, polycystic ovary syndrome, and fertility. Curr Opin Endocrinol Diabetes Obes. 2022;29(6):554–559.

22. Sun M, Li J, Wang Y, Wei Y, Li Y, Cui L, et al. Proteomic analysis of human ovarian granulosa cells in women with polycystic ovary syndrome using data-independent acquisition mass spectrometry. Front Endocrinol. 2022;13:833871.

23. Zhou HJ, Wang LQ, Wu X, Yang HX, Zhang XL, Hua KQ. Weighted gene co-expression network analysis reveals potential biomarkers for granulosa cell apoptosis in polycystic ovary syndrome. Front Genet. 2022;13:867997.

24. Choi SW, Friso S. Modulation of DNA methylation by one-carbon metabolism: a milestone for healthy aging. Nutr Res Pract. 2023;17(4):597–615.

25. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks. Nucleic Acids Res. 2021;49(D1):D605–D612.

26. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol. 2017;18(1):83.

27. Razzaq A, Strogatz C, Guertin S, et al. Recent web platforms for multi-omics integration unlocking biological complexity. Appl Sci. 2025;15(1):329.

28. Subramanian I, Verma S, Kumar S, Jain A, Anamika K. Multi-omics data integration, interpretation, and its application. Bioinform Biol Insights. 2020;14:1177932219899051.

29. Parker J, O'Brien C, Hawrelak J, Gersh FL. Polycystic ovary syndrome: an evolutionary adaptation to lifestyle and the environment. Int J Environ Res Public Health. 2022;19(3):1336.

30. Stener-Victorin E, Deng Q, Couto Alves A, et al. Polycystic ovary syndrome. Nat Rev Dis Primers. 2024;10(1):27.

31. Feng W, Zhang Y, Pan Y, et al. Association of three missense mutations in homocysteine-related MTHFR and MTRR gene with risk of polycystic ovary syndrome in Southern Chinese women. Reprod Biol Endocrinol. 2021;19(1):5.

32. Rawat K, Sandhu A, Gautam V, Saha PK, Saha L. Role of genomic DNA methylation in PCOS pathogenesis: a systematic review and meta-analysis. Mol Hum Reprod. 2022;28(7):gaac024.

33. Menezo Y, Clement P, Clement A, Elder K. Methylation: an ineluctable biochemical and physiological process essential to the transmission of life. Int J Mol Sci. 2020;21(23):9311.

34. Gong C, Chen Y, Shan A, et al. The role of one-carbon metabolism and methyl donors in medically assisted reproduction. Nutrients. 2024;16(9):1339.

35. Carrasco-Cabezas M, Assmann TS, Martínez P, et al. Folate and vitamin B12 levels in Chilean women with PCOS and their association with metabolic outcomes. Nutrients. 2024;16(12):1937.

36. Muñoz-Gómez M, Quesada-Varas M, Aparicio-Marín E, García-Bolufer L, Cano R, López-García AI. Prevalence of MTHFR polymorphisms and metabolic outcomes in Chilean women with polycystic ovary syndrome. Endocr Abstr. 2023;90:EP1157.

37. Zhu X, Hong X, Chen L, Xuan Y, Huang K, Wang B. Association of MTHFR C677T and A1298C polymorphisms with genetic susceptibility to polycystic ovary syndrome: a PRISMA-compliant meta-analysis. Gene. 2019;710:232–239.

38. Abd-alkareem MA, Omear HA. INSR and PLIN polymorphism in women with polycystic ovary syndrome and its correlation with insulin resistance. Int J Med Sci. 2022;3(3):93–110.

39. Parker J, Nekoonam S, Hosseini B, et al. Recognizing the role of insulin resistance in polycystic ovary syndrome: a paradigm shift. J Clin Med. 2025;14(12):4021.

40. Baysoy A, Bai Z, Satija R, Fan R. The technological landscape and applications of single-cell multi-omics. Nat Rev Mol Cell Biol. 2023;24(10):695–713.

41. Medrano RF, de Oliveira CA. A cost-effective and rapid PCR-RFLP method for genotyping MTHFR C677T. Genomics Data. 2014;2:234–240.

42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

43. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–140.

44. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. Bioinformatics. 2011;27(11):1571–1572.

45. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363–1369.

46. Mo Q, Wang S, Seshan VE, et al. Pattern discovery and cancer gene identification in integrated cancer genomic data. Proc Natl Acad Sci USA. 2013;110(11):4245–4250.