

Identification of key molecular targets in nicotine-induced spontaneous abortion through network toxicology and multi-omics integration

Kaiyan Yang^{1*}, Xinying Song^{1*}, Pengyan Zhai¹, Huiyan Wang¹, Wenbo Zhou¹

ABSTRACT

INTRODUCTION Spontaneous abortion (SA) remains a prevalent reproductive health challenge, with tobacco-derived nicotine emerging as a significant risk factor. This study sought to decipher the molecular underpinnings of nicotine-induced pregnancy loss through comprehensive multi-omics profiling to identify novel biomarkers and potential intervention targets.

METHODS An integrated bioinformatics approach was implemented combining network toxicology, transcriptomic profiling, and machine learning algorithms to elucidate nicotine's pathological mechanisms in SA. The analytical pipeline encompassed functional enrichment analysis, expression pattern characterization, diagnostic biomarker evaluation, immune microenvironment assessment, causal inference through Mendelian randomization (MR), single-cell RNA sequencing, and molecular docking simulations of nicotine-protein interactions.

RESULTS Differential expression analysis identified 2451 DEGs between SA and control groups. Integration with 645 nicotine targets yielded 69 candidate genes. Five hub genes – CTNNB1, TINAGL1, RAD50, FGF2, and VIM – were identified as critical mediators in nicotine-triggered SA pathogenesis. These genes exhibited good diagnostic capability and significant correlations with immune cell infiltration patterns. MR analysis confirmed causal associations for RAD50 (OR=1.08; 95% CI: 1.01–1.15, $p=0.019$) and FGF2 (OR=1.21; 95% CI: 1.11–1.33, $p=4.69\times10^{-5}$). Single-cell analyses revealed cell-type-specific expression signatures across trophoblasts and immune populations, while molecular docking showed binding energies ranging from -5.0 to -6.9 kcal/mol.

CONCLUSIONS This study illuminates a complex, multi-target mechanism underlying tobacco-induced pregnancy loss. This work advances understanding of nicotine's reproductive toxicity pathways and identifies promising molecular targets for early diagnosis and further mechanistic studies.

AFFILIATION

¹ Changzhou Maternal and Child Health Care Hospital, Changzhou Medical Center, Nanjing Medical University, Changzhou, China

*Contributed equally

CORRESPONDENCE TO

Wenbo Zhou. Medical Research Center, Changzhou Maternal and Child Health Care Hospital, Changzhou Medical Center, Nanjing Medical University, No. 16 Dingxiang Road, Changzhou, Jiangsu Province, Zip Code 213000, China

E-mail: wenbozhou@njmu.edu.cn

ORCID iD: <https://orcid.org/0000-0001-7254-3731>

KEYWORDS

nicotine, spontaneous abortion, network toxicology, multi-omics, molecular docking

Received: 30 January 2026

Revised: 7 August 2026

Accepted: 14 August 2026

INTRODUCTION

Spontaneous abortion (SA) refers to pregnancy loss occurring before 20 weeks of gestation or when the fetal weight is <500 g, and it is one of the most common complications during early pregnancy¹. According to recent epidemiological surveys, approximately 15–20% of clinically recognized pregnancies end in SA; when including undetected biochemical pregnancies, the overall incidence may reach as high as 30–50%¹. SA not only leads to severe physiological complications such as hemorrhage and infection in women but also frequently

triggers psychological disorders such as anxiety and depression, imposing substantial physical, emotional, and economic burdens on individuals, families, and society². Although more than half of early SA can be attributed to genetic abnormalities such as embryonic chromosomal aneuploidy, nearly half of cases remain idiopathic, suggesting a significant role of modifiable factors such as environmental exposures and lifestyle in the etiology of SA³.

Among various modifiable environmental factors, tobacco exposure is widely recognized as one of the most prominent risk factors affecting reproductive health in women of reproductive age⁴. Numerous prospective cohort studies and case-control studies have consistently confirmed a significant association between active or passive smoking during pregnancy and increased risk of spontaneous abortion⁵.

Nicotine, the most abundant addictive alkaloid in tobacco, is considered the primary bioactive component mediating smoking-related reproductive toxicity. Studies have shown that nicotine can rapidly cross the placental barrier and enter the fetal circulation, with concentrations in amniotic fluid, fetal blood, and tissues even exceeding those in maternal plasma⁶. Known mechanisms by which nicotine contributes to miscarriage involve multiple biological levels: inducing vasoconstriction of uterine and placental arteries via activation of the sympathetic nervous system, leading to impaired placental perfusion and fetal hypoxia; triggering oxidative stress and endoplasmic reticulum stress in placental trophoblast cells, thereby compromising placental barrier integrity; and modulating decidual immune cell function, disrupting maternal-fetal immune tolerance⁷.

Although these studies provide important biological evidence for nicotine's reproductive toxicity, significant knowledge gaps remain. First, the precise molecular targets and key signaling pathways through which nicotine acts at the maternal-fetal interface have not been fully elucidated. Second, the differential sensitivity of various cell types to nicotine and their specific regulatory networks remain unclear. Network toxicology, an emerging interdisciplinary field integrating systems biology and toxicology, enables the systematic elucidation of the toxicity mechanisms of exogenous chemicals by

constructing multidimensional interaction networks among 'compounds, targets, and diseases'⁸. Mendelian randomization (MR) analysis, an innovative epidemiological method, uses genetic variants as instrumental variables to infer causal relationships between exposures and disease outcomes, effectively minimizing confounding and reverse causality⁹. The rapid advancement of single-cell RNA sequencing (scRNA-seq) technology has provided a revolutionary tool for dissecting cellular heterogeneity and molecular mechanisms in pregnancy-related conditions¹⁰. Molecular docking, based on structural biology and computational chemistry, simulates the three-dimensional binding modes between small molecule ligands and target proteins, quantitatively evaluating the affinity and specificity of molecular interactions¹¹.

This study establishes a systematic research framework integrating network toxicology prediction, transcriptomic differential analysis, single-cell sequencing validation, Mendelian randomization for causal inference, and molecular docking simulation to systematically identify key molecular targets of nicotine associated with spontaneous abortion and elucidate their expression profiles and regulatory networks across different cell types at the maternal-fetal interface.

METHODS

Identification of potential targets for nicotine

We systematically identified potential targets of nicotine by integrating multiple authoritative drug target prediction databases: 1) SwissTargetPrediction database with species set to 'Homo sapiens' and predicted targets with probability values ≥ 0.1 ; 2) STITCH compound-protein interaction database using a confidence score threshold of ≥ 0.4 ; 3) DGIdb drug-gene interaction database to collect reported nicotine interaction targets; and 4) Comparative Toxicogenomics Database screened for interaction targets supported by at least one peer-reviewed publication.

Transcriptomic differential expression analysis and biomarker screening

We analyzed the endometrial tissue transcriptomic dataset GSE165004 from the GEO database,

comprising decidual tissue samples from 24 recurrent SA patients and 24 normal early pregnancy women. After quality control, we performed normalization and differential expression analysis using the *limma* package in R. Differentially expressed genes (DEGs) were screened using adjusted $p < 0.05$. Subsequently, intersection analysis was performed between DEGs and the nicotine target gene set.

Functional enrichment and pathway analysis

Multi-dimensional functional annotation and pathway enrichment analyses were performed on the intersected gene set to elucidate the underlying molecular mechanisms. Gene Ontology (GO) analysis systematically examined gene functional hierarchies from three perspectives: biological process, molecular function, and cellular component. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis focused on revealing signal transduction networks and metabolic pathways involving these genes. All enrichment analyses were implemented using the *clusterProfiler* R package¹², with significance levels set at $p < 0.05$ (FDR corrected).

Machine learning algorithms for key gene identification

To identify core genes with the highest discriminative power from high-dimensional gene lists, this study employed two complementary machine learning feature selection methods: LASSO regression and Support Vector Machine-Recursive Feature Elimination (SVM-RFE). LASSO regression analysis was implemented using the *glmnet* R package¹³, constructing an L1-regularized logistic regression model with group labels as response variables and intersection gene expression profiles as predictor variables. The optimal penalty parameter λ (lambda.min) was determined through 10-fold cross-validation, and genes corresponding to non-zero coefficients were retained as first-tier screening results. SVM-RFE analysis was performed using the *e1071* and *caret* R packages¹⁴, constructing support vector machine classifiers and iteratively removing features with the smallest weights until the remaining feature set achieved minimal classification error. Finally, the intersection of genes from both methods was taken as the final core signature gene set to enhance the

robustness and generalizability of the screening results.

Expression profiling and diagnostic evaluation of key genes

Expression level visualization and diagnostic efficacy analysis were performed on the screened core genes within the GSE165004 dataset. Box plots were generated using the *ggplot2* package to intuitively display expression difference patterns of each gene between recurrent SA and NEP groups. Simultaneously, receiver operating characteristic (ROC) curves and their corresponding areas under the curve (AUC) were calculated for each core gene to quantitatively assess their diagnostic potential as biomarkers.

Immune infiltration analysis

The CIBERSORTx algorithm was employed to perform deconvolution analysis on gene expression profiles from GSE165004, estimating the relative abundance of 22 immune cell subsets¹⁵. The analysis used limma-normalized expression matrices as input data, with Leukocyte Gene Signature Matrix (LM22) selected as the reference signature matrix. Through Spearman rank correlation analysis, we systematically calculated correlation coefficients between core gene expression levels and the abundance of various immune cells, and visualized the gene-immune cell association network in heatmap format. Wilcoxon rank-sum tests were used to compare differences in immune cell composition between groups to reveal potential immune imbalance mechanisms.

MR analysis

To assess the potential causal association between identified key target genes and SA, we employed a two-sample Mendelian randomization (MR) design. The validity of this approach relies on three core assumptions: 1) the relevance assumption, requiring that the genetic instruments are strongly associated with the exposure; 2) the independence assumption, requiring that the instruments are not associated with confounders of the exposure-outcome relationship; and 3) the exclusion restriction assumption, requiring that the instruments influence the outcome only through the exposure, not via alternative pathways.

To satisfy the relevance assumption, instrumental variables (IVs) were selected from expression quantitative trait loci (eQTL) data at a genome-wide significance threshold ($p < 5 \times 10^{-8}$). Only single nucleotide polymorphisms (SNPs) with F-statistic > 10 were retained, ensuring that the instruments were sufficiently strong and effectively minimizing weak instrument bias. When a selected IV was not present in the outcome data, it was excluded from the MR analysis. For the independence assumption, we obtained eQTL summary statistics from the IEU Open GWAS resource, which derives primarily from the eQTLGen consortium ($n = 31684$ individuals)¹⁶, while outcome data for SA were drawn from the FinnGen R12 dataset (23167 cases and 199279 controls), an independent Finnish population cohort¹⁷ ([Supplementary file](#) Table 1). Both the exposure and outcome GWAS were derived from European-ancestry populations and are independent, with no sample overlap. Furthermore, linkage disequilibrium (LD) clumping was performed ($r^2 < 0.1$, window size 10000 kb) to ensure that the selected SNPs were independent of one another. Target genes for which no eQTLs were available in the eQTLGen resource, or for which no instrumental variable passed the selection criteria, were excluded from causal estimation.

The primary causal effect was estimated using the inverse variance weighted (IVW) method. To thoroughly assess the exclusion restriction assumption, multiple supplementary analyses were implemented to detect and adjust for potential horizontal pleiotropy. Specifically, we applied the weighted median and MR-Egger regression methods, which are more robust to pleiotropy than the IVW approach under different scenarios. Heterogeneity among the IVs was evaluated with Cochran's Q test, while the MR-Egger intercept test was used as a formal test of directional pleiotropy. In addition, the MR-PRESSO global test was employed to identify outlier SNPs that might violate the exclusion restriction. Leave-one-out analyses were further conducted to confirm that the overall causal estimates were not driven by any single variant.

Single-cell validation of core target gene expression

To validate the expression characteristics of the aforementioned core target genes in recurrent

abortion at single-cell resolution, we analyzed the single-cell transcriptomic dataset GSE214607 from the GEO database, which encompassed decidual tissue samples from 3 SA patients and 5 control women. The original gene-cell expression matrix was generated using the 10x Genomics platform and processed using the *Seurat* package in R. Quality control was first performed on cells, and potential doublets were removed using *DoubletFinder*. After using *LogNormalize*, highly variable genes were selected and the *Harmony* algorithm was employed to correct for sample batch effects. Based on t-SNE dimensionality reduction clustering, cell type annotation was performed according to marker genes from reference literature. Subsequently, the core target genes identified in the previous study were mapped to various cell types, and their expression distribution characteristics were visualized through dot plots and feature plots. Wilcoxon rank-sum tests were applied to compare expression differences of target genes in each cell type between SA and control groups, with results displayed through violin plots.

Molecular docking validation

To predict the binding capacity between nicotine and core target proteins, this study employed the CB-Dock2 online server for molecular docking analysis. CB-Dock2 integrates the curvature-driven binding site identification algorithm with the AutoDock Vina docking engine, enabling efficient prediction of receptor protein binding cavities and ligand docking simulation¹⁸. Three-dimensional structures of core target proteins were obtained from the Protein Data Bank. For proteins lacking experimentally resolved structures, high-precision structural models predicted by AlphaFold2 were employed. The three-dimensional structure of nicotine was retrieved from the PubChem database (CID: 89594). Docking parameters were set to program defaults, binding energy scoring employed the Vina scoring function, and the highest affinity conformation and predicted binding free energy (kcal/mol) for each complex were systematically recorded.

Statistical analysis

All bioinformatics and statistical analyses were performed using R software (version 4.3.2).

Differential expression analysis was conducted with the *limma* package, applying the Benjamini-Hochberg procedure for multiple testing correction (adjusted $p < 0.05$). Diagnostic performance was evaluated by receiver operating characteristic (ROC) curves, with an area under the curve (AUC) > 0.7 considered indicative of acceptable discriminatory ability. For immune infiltration comparisons, Wilcoxon rank-sum tests were used for between-group comparisons, and Spearman's rank correlation was used for association analyses. MR analyses were performed using the *TwoSampleMR* (version 0.6.8) and *MR-PRESSO* (version 1.0) packages, the IVW method served as the primary estimator, with a $p < 0.05$ considered statistically significant. Heterogeneity was evaluated with Cochran's Q test, and horizontal pleiotropy was assessed using the MR-Egger intercept test; for both tests, $p < 0.05$ was interpreted as evidence of significant heterogeneity or pleiotropy, respectively. Molecular docking was carried out using CB-Dock2, and binding energies ≤ -5.0 kcal/mol were considered indicative of significant binding affinity.

RESULTS

Network toxicology prediction of key targets in nicotine-induced SA

Integration of data from SwissTargetPrediction, STITCH, DGIdb, and CTD databases yielded 645 potential nicotine targets ([Supplementary file](#) Figure 1). Differential expression analysis of GEO datasets revealed 2451 DEGs, visualized through volcano plots and heatmaps ([Supplementary file](#) Figures 2 and 3). We identified 69 overlapping genes between nicotine targets and SA-related genes ([Supplementary file](#) Figure 4), which were used to construct a protein-protein interaction network ([Supplementary file](#) Figure 5).

GO enrichment analysis revealed that these targets were primarily involved in hormone response (glucocorticoid and steroid), apoptosis regulation, and mechanical stimulus response. Cellular components were enriched in outer membranes, membrane rafts, and synaptic structures, while molecular functions concentrated on ubiquitin ligase activity, phosphatase activity, and protein binding ([Supplementary file](#) Figure 6). KEGG pathway analysis demonstrated significant enrichment in viral infection pathways,

cancer-related pathways, apoptotic pathways, and the PD-L1 pathway ([Supplementary file](#) Figure 7).

Machine learning identification of key diagnostic genes

Using the 69 candidate target genes, we employed complementary machine learning approaches for feature selection. LASSO regression identified 15 genes with non-zero coefficients ([Supplementary file](#) Figures 8 and 9), while SVM-RFE algorithm indicated an optimal feature number of 5 ([Supplementary file](#) Figure 10). Intersection analysis revealed five key genes: CTNNB1, FGF2, RAD50, TINAGL1, and VIM ([Supplementary file](#) Figure 11).

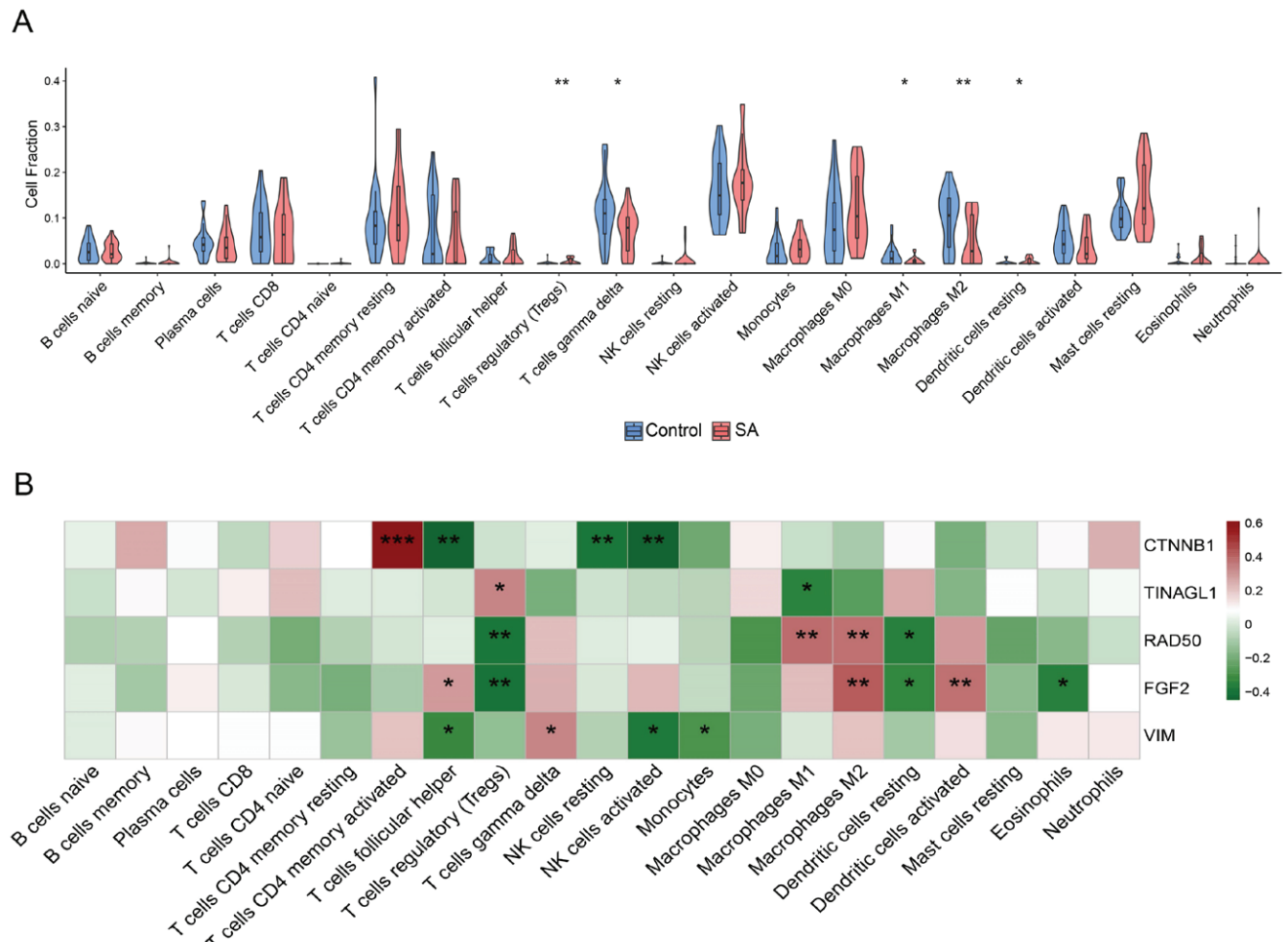
Expression analysis demonstrated significant differences between SA and control groups for all five genes: TINAGL1 was upregulated in SA, while CTNNB1, FGF2, RAD50, and VIM were downregulated ([Supplementary file](#) Figure 12). ROC curve analysis validated the diagnostic efficacy of all five genes, with AUC values ranging from 0.747 to 0.969 ([Supplementary file](#) Figure 13).

Immune infiltration analysis

The distribution of 22 immune cell types revealed distinct compositional patterns between groups ([Supplementary file](#) Figure 14), with activated mast cells showing no infiltration signal across all samples. Correlation analysis revealed complex intercellular interactions, such as the positive correlation between M2 macrophages and plasma cells and the associations among other immune subsets ([Supplementary file](#) Figure 15). The SA group exhibited significant alterations in regulatory T cells, $\gamma\delta$ T cells, macrophages (M1 and M2), and dendritic cell proportions (Figure 1A).

Correlation analysis between key genes and immune cell infiltration showed that CTNNB1 positively correlated with activated memory CD4+ T cells while negatively correlating with follicular helper T cells and NK cells. RAD50 significantly correlated with regulatory T cells, macrophages, and dendritic cells. FGF2 positively correlated with M2 macrophages and activated dendritic cells but negatively with regulatory T cells. Both TINAGL1 and VIM were associated with specific T cell subsets and myeloid cells (Figure 1B).

Figure 1. Immune infiltration analysis: A) Violin plots comparing the distributions of each immune-cell subset between control and SA samples; B) Spearman correlation heatmap between the infiltration levels of the immune-cell subsets and the five hub genes



SA: spontaneous abortion. *p<0.05, **p<0.01, ***p<0.001.

MR analysis validates causal associations

Since no expression-associated signals for TINAGL1 were found in the eQTLGen database, we employed two-sample MR to evaluate causal associations between four potential target genes (CTNNB1, RAD50, FGF2, and VIM) and SA. The IVs are shown in [Supplementary file](#) Table 2, with the F-statistic ranging from 30.58 to 1140.08, indicating no evidence of weak instrument bias.

The MR analysis results showed significant causal associations between FGF2 (95% CI: 1.11–1.33, $p=4.69\times10^{-5}$) and RAD50 (OR=1.08; 95% CI: 1.01–1.15, $p=0.019$) and SA risk ([Supplementary file](#): Table 3 and Figure 16). Sensitivity analyses further

supported the robustness of these findings (MR-Egger intercept $p>0.05$) ([Supplementary file](#) Tables 3 and 4).

Single-cell transcriptome analysis validates cell-specific expression of core genes

After quality control and standardization, we obtained 48051 high-quality cells, including 20799 cells from the SA group and 27252 cells from the control group. Through t-SNE dimensionality reduction combined with cell type marker genes indicated by the data source references¹⁹, we identified 11 cell types consistent with those reported in the literature, including B cells, T cells, dendritic cells (DCs),

decidual NK cells (dNKs), decidual macrophages (dM), monocytes (Mon), neutrophils, decidual stromal cells (DSCs), endothelial cells (Endo), epithelial cells (Epi), and extravillous trophoblasts (EVT).

Dot plot and feature plot analyses revealed distinct cell type-specific expression patterns of the five core genes ([Supplementary file](#) Figures 18 and 19). VIM showed high expression in stromal-derived and immune cells, particularly in DSCs, dM, and DCs. CTNNB1 was enriched in tissue barrier-related cells (Endo, DSCs, EVT) and specific immune subsets. TINAGL1 exhibited highly tissue-specific expression, primarily in EVT and Endo cells. FGF2 maintained low basal expression across most cell types, with high expression only in DSCs. RAD50 showed relatively uniform expression across all cell populations, with slight enrichment in EVT and DSCs.

Differential expression analysis revealed significant cell type-specific changes ([Supplementary file](#) Figure 20). TINAGL1 was significantly upregulated across multiple cell types (Epi, DSCs, dNK, dM, and EVT). CTNNB1 was significantly downregulated in EVT, Mon, dNK cells, and DSCs, but upregulated in Epi. RAD50 was significantly downregulated in DSCs, Mon, and Endo, but upregulated in EVT and Epi. FGF2 was significantly downregulated in DSCs, dM, EVT, and Mon, showing upregulation trends only in T cells. VIM was significantly upregulated across multiple cell types including Epi, dNK, dM, DCs, T cells, and neutrophils.

Molecular docking analysis

Crystal structures of four target proteins were obtained from the PDB: β -catenin (CTNNB1, ID: 1JDH), RAD50 (ID: 5GOX), FGF2 (ID: 1CVS), and VIM (ID: 3KLT). For TINAGL1, we utilized its AlphaFold structure (ID: AF-Q9GZM7-F1).

Nicotine formed stable ligand-protein complexes with all five target proteins ([Supplementary file](#) Figure 21), with binding free energies ranging from -5.0 to -6.9 kcal/mol, consistent with moderate to high-affinity small molecule-protein interactions. Notably, nicotine formed the most stable complex with cytoskeletal protein VIM (-6.9 kcal/mol).

DISCUSSION

This study employed an integrated multi-omics

approach to systematically elucidate the molecular mechanisms by which nicotine may contribute to SA through multi-target, multi-pathway, and cell-specific regulatory networks. Through the integration of network pharmacology, transcriptomics, single-cell genomics, and molecular docking techniques, this study identified five core targets and demonstrated their aberrant expression in SA and associations with immune microenvironmental alterations, establishing a novel framework for understanding nicotine's pathological effects on pregnancy maintenance.

Network pharmacology analysis revealed that common targets between nicotine and SA were primarily enriched in hormonal response regulation, apoptotic pathways, and immune checkpoint processes. These findings align with previous studies demonstrating nicotine's effects on hormonal balance, cellular apoptosis, and immune microenvironment modulation during pregnancy²⁰. The enrichment of viral infection-related pathways is particularly intriguing, as emerging evidence suggests that disruption of antiviral immune pathways may compromise placental immune surveillance mechanisms²¹. Recent research has demonstrated that placental tissues express numerous antiviral response genes that not only protect against viral pathogens but also regulate trophoblast invasion and differentiation, suggesting potential convergence between immune defense mechanisms and pregnancy maintenance pathways that may be disrupted by nicotine exposure²².

The five core target genes identified through integrated machine learning approaches all demonstrated substantial diagnostic value with clear biological plausibility for involvement in pregnancy maintenance. CTNNB1, encoding β -catenin as the core transducer of Wnt signaling, plays a pivotal role in embryonic implantation and placental formation²³. The single-cell analysis revealed CTNNB1 expression across multiple placental cell types, particularly abundant in Endo, DSCs, and EVT, but significantly downregulated in DSCs, dNK, and EVT in abortion samples. This observation aligns with previous studies demonstrating associations between aberrant Wnt/ β -catenin signaling and increased abortion risk^{23,24}.

TINAGL1 exhibited remarkable diagnostic value, with significantly elevated expression in decidual tissues from abortion patients. As an extracellular

matrix protein involved in cellular adhesion, migration, and angiogenesis²⁵, its role as a positive regulator of angiogenesis and expression in mouse uterine decidua following implantation is particularly relevant, considering that insufficient trophoblast invasion constitutes an important pathological basis for abortion. Single-cell analysis demonstrated that TINAGL1 is primarily expressed in EVT and endothelial cells, both crucial for placental vascular remodeling. Previous studies have shown that TINAGL1 functions as an integrin ligand involved in endometrial stromal cell adhesion during embryo implantation²⁶ and co-localizes with extracellular matrix proteins during pre-implantation embryonic development²⁷. It is hypothesized that TINAGL1 upregulation in abortion patients may represent a pathological adaptive response, potentially reflecting a compensatory mechanism attempting to repair damaged maternal-fetal interfaces, though this excessive response may ultimately disrupt normal extracellular matrix balance.

FGF2 demonstrated strong causal evidence for association with SA in MR analysis, aligning with previous studies showing that nicotine exposure alters FGF2 gene and protein expression and impair FGF2 signaling systems²⁸. As a crucial regulator of angiogenesis and tissue repair, FGF2 promotes placental vascular network formation and decidualization during normal pregnancy²⁹. The single-cell analysis revealed significant FGF2 downregulation in decidual stromal cells from abortion patients, consistent with reduced FGF2 expression observed in placental tissues from women with recurrent pregnancy loss. The significant positive correlation between FGF2 and M2 macrophage abundance supports its involvement in maternal-fetal interface immune regulation, aligning with previous reports of FGF2's role in promoting M2 macrophage polarization³⁰. Given that M2 macrophages play protective roles in maintaining maternal-fetal immune tolerance³¹, these findings provide new perspectives for understanding how environmental factors like smoking affect pregnancy outcomes through specific molecular pathways.

RAD50, a core component of the DNA double-strand break repair complex³², showed complex regulatory patterns in the analyses. While MR analysis suggested

that increased RAD50 in blood correlates positively with abortion risk, the transcriptomic data revealed significant downregulation in abortion decidual tissue, with single-cell analysis uncovering complex cell type-specific expression patterns. Given that smoking causes DNA damage³³, and that DNA repair capacity is crucial during early embryonic development, it is hypothesized that nicotine may interfere with RAD50-mediated DNA repair function, leading to increased genomic instability in placental cells. The correlation between RAD50 and regulatory T cells/macrophages abundance is consistent with emerging research proposing crossover mechanisms between DNA damage response and immune regulation³⁴.

VIM demonstrated the highest binding affinity with nicotine in molecular docking analysis, suggesting it may represent an important direct target. As an intermediate filament protein involved in cytoskeletal organization and epithelial-mesenchymal transition (EMT) processes³⁵, VIM showed significant upregulation across multiple immune cell types in SA patients, consistent with its functions in immune cell activation³⁶.

Limitations

While the integration of network toxicology with multi-omics data provides novel insights into environmentally induced pregnancy loss mechanisms, several methodological limitations warrant consideration. These include: the necessity for experimental validation of predicted nicotine-protein interactions; limited representativeness due to restricted decidual tissue samples for single-cell sequencing; interpretive constraints of the MR analyses, which reflect cumulative rather than stage-specific genetic effects, and the use of data from predominantly European participants, potentially limiting the generalizability of the findings to other populations; the potential for residual confounding in the transcriptomic and immune infiltration analyses, as they are based on observational data; and the inherent limitations of bulk transcriptomics in capturing cell-specific alterations. These constraints highlight critical directions for future investigations.

CONCLUSIONS

Through network toxicology and multi-omics

integration, this study elucidated the molecular mechanisms underlying nicotine-induced spontaneous abortion (SA) and identified several key target genes. These findings advance the understanding of how environmental exposures disrupt maternal-fetal homeostasis and provide a molecular framework for biomarker-based risk prediction and further mechanistic studies.

REFERENCES

1. Quenby S, Gallos ID, Dhillon-Smith RK, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. *Lancet*. 2021;397(10285):1658-1667. doi:[10.1016/S0140-6736\(21\)00682-6](https://doi.org/10.1016/S0140-6736(21)00682-6)
2. Farren J, Mitchell-Jones N, Verbakel JY, Timmerman D, Jalmbrant M, Bourne T. The psychological impact of early pregnancy loss. *Hum Reprod Update*. 2018;24(6):731-749. doi:[10.1093/humupd/dmy025](https://doi.org/10.1093/humupd/dmy025)
3. Dimitriadis E, Menkhorst E, Saito S, Kutteh WH, Brosens JJ. Recurrent pregnancy loss. *Nat Rev Dis Primers*. 2020;6(1):98. doi:[10.1038/s41572-020-00228-z](https://doi.org/10.1038/s41572-020-00228-z)
4. Dechanet C, Anahory T, Mathieu Daude JC, et al. Effects of cigarette smoking on reproduction. *Hum Reprod Update*. 2011;17(1):76-95. doi:[10.1093/humupd/dmq033](https://doi.org/10.1093/humupd/dmq033)
5. Pineles BL, Park E, Samet JM. Systematic review and meta-analysis of miscarriage and maternal exposure to tobacco smoke during pregnancy. *Am J Epidemiol*. 2014;179(7):807-823. doi:[10.1093/aje/kwt334](https://doi.org/10.1093/aje/kwt334)
6. Luck W, Nau H, Hansen R, Steldinger R. Extent of nicotine and cotinine transfer to the human fetus, placenta and amniotic fluid of smoking mothers. *Dev Pharmacol Ther*. 1985;8(6):384-395. doi:[10.1159/000457063](https://doi.org/10.1159/000457063)
7. Wong MK, Holloway AC, Hardy DB. Nicotine directly induces endoplasmic reticulum stress response in rat placental trophoblast giant cells. *Toxicol Sci*. 2016;151(1):23-34. doi:[10.1093/toxsci/kfw019](https://doi.org/10.1093/toxsci/kfw019)
8. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*. 2008;4(11):682-690. doi:[10.1038/nchembio.118](https://doi.org/10.1038/nchembio.118)
9. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89-R98. doi:[10.1093/hmg/ddu328](https://doi.org/10.1093/hmg/ddu328)
10. Trapnell C. Defining cell types and states with single-cell genomics. *Genome Res*. 2015;25(10):1491-1498. doi:[10.1101/gr.190595.115](https://doi.org/10.1101/gr.190595.115)
11. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*. 2004;3(11):935-949. doi:[10.1038/nrd1549](https://doi.org/10.1038/nrd1549)
12. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-287. doi:[10.1089/omi.2011.0118](https://doi.org/10.1089/omi.2011.0118)
13. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1-22. doi:[10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01)
14. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics*. 2018;15(1):41-51. doi:[10.21873/cgp.20063](https://doi.org/10.21873/cgp.20063)
15. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457. doi:[10.1038/nmeth.3337](https://doi.org/10.1038/nmeth.3337)
16. Vösa U, Claringbould A, Westra HJ, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet*. 2021;53(9):1300-1310. doi:[10.1038/s41588-021-00913-z](https://doi.org/10.1038/s41588-021-00913-z)
17. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613(7944):508-518. doi:[10.1038/s41586-022-05473-8](https://doi.org/10.1038/s41586-022-05473-8)
18. Liu Y, Yang X, Gan J, Chen S, Xiao ZX, Cao Y. CB-Dock2: improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res*. 2022;50(W1):W159-W164. doi:[10.1093/nar/gkac394](https://doi.org/10.1093/nar/gkac394)
19. Wei P, Dong M, Bi Y, et al. Identification and validation of a signature based on macrophage cell marker genes to predict recurrent miscarriage by integrated analysis of single-cell and bulk RNA-sequencing. *Front Immunol*. 2022;13:1053819. doi:[10.3389/fimmu.2022.1053819](https://doi.org/10.3389/fimmu.2022.1053819)
20. Zdravkovic T, Genbacev O, McMaster MT, Fisher SJ. The adverse effects of maternal smoking on the human placenta: a review. *Placenta*. 2005;26(suppl A):S81-S86. doi:[10.1016/j.placenta.2005.02.003](https://doi.org/10.1016/j.placenta.2005.02.003)
21. Delorme-Axford E, Sadovsky Y, Coyne CB. The placenta as a barrier to viral infections. *Annu Rev Virol*. 2014;1(1):133-146. doi:[10.1146/annurev-virology-031413-085524](https://doi.org/10.1146/annurev-virology-031413-085524)
22. Cornish EF, Filipovic I, Åsenius F, Williams DJ, McDonnell T. Innate immune responses to acute viral infection during pregnancy. *Front Immunol*. 2020;11:572567. doi:[10.3389/fimmu.2020.572567](https://doi.org/10.3389/fimmu.2020.572567)
23. Li S, Li N, Zhu P, Wang Y, Tian Y, Wang X. Decreased β -catenin expression in first-trimester villi and decidua of patients with recurrent spontaneous abortion. *J Obstet Gynaecol Res*. 2015;41(6):904-911. doi:[10.1111/jog.12647](https://doi.org/10.1111/jog.12647)
24. Zhou M, Gao Y, Wu S, et al. Ubiquitin-specific protease 7 prevents recurrent spontaneous abortion by targeting enhancer of zeste homolog 2 to regulate trophoblast proliferation, apoptosis, migration, and invasion through the Wnt/ β -catenin pathway. *Biol Reprod*. 2023;109(2):204-214. doi:[10.1093/biolre/ioad053](https://doi.org/10.1093/biolre/ioad053)
25. Gu S, Zang X, Jiang L, et al. Differential microRNA expression in porcine endometrium related to spontaneous embryo loss during early pregnancy. *Int J Mol Sci*. 2022;23(15):8157. doi:[10.3390/ijms23158157](https://doi.org/10.3390/ijms23158157)

26. Tajiri Y, Igarashi T, Li D, et al. Tubulointerstitial nephritis antigen-like 1 is expressed in the uterus and binds with integrins in decidualized endometrium during postimplantation in mice. *Biol Reprod.* 2010;82(2):263-270. doi:[10.1095/biolreprod.109.080028](https://doi.org/10.1095/biolreprod.109.080028)
27. Sakurai M, Sato Y, Mukai K, et al. Distribution of tubulointerstitial nephritis antigen-like 1 and structural matrix proteins in mouse embryos during preimplantation development in vivo and in vitro. *Zygote.* 2014;22(2):259-265. doi:[10.1017/S0967199412000469](https://doi.org/10.1017/S0967199412000469)
28. Zhang WX, Chen HJ, Fan J, et al. The association between maternal nicotine exposure and adipose angiogenesis in female rat offspring: a mechanism of adipose tissue function changes. *Toxicol Lett.* 2020;318:12-21. doi:[10.1016/j.toxlet.2019.10.007](https://doi.org/10.1016/j.toxlet.2019.10.007)
29. Zhang X, Mo Y, Lu C, Su Z, Li X. Fibroblast growth factors and endometrial decidualization: models, mechanisms, and related pathologies. *J Zhejiang Univ Sci B.* 2025;26(6):573-588. doi:[10.1631/jzus.B2300830](https://doi.org/10.1631/jzus.B2300830)
30. Yi L, Chen Y, Zhang Y, et al. Deleting fibroblast growth factor 2 in macrophages aggravates septic acute lung injury by increasing M1 polarization and inflammatory cytokine secretion. *Mol Biomed.* 2024;5(1):50. doi:[10.1186/s43556-024-00203-0](https://doi.org/10.1186/s43556-024-00203-0)
31. Houser BL, Tilburgs T, Hill J, Nicotra ML, Strominger JL. Two unique human decidual macrophage populations. *J Immunol.* 2011;186(4):2633-2642. doi:[10.4049/jimmunol.1003153](https://doi.org/10.4049/jimmunol.1003153)
32. Hopfner KP. Mre11-Rad50: the DNA end game. *Biochem Soc Trans.* 2023;51(2):527-538. doi:[10.1042/BST20220754](https://doi.org/10.1042/BST20220754)
33. Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene.* 2002;21(48):7435-7451. doi:[10.1038/sj.onc.1205803](https://doi.org/10.1038/sj.onc.1205803)
34. Nakad R, Schumacher B. DNA damage response and immune defense: links and mechanisms. *Front Genet.* 2016;7:147. doi:[10.3389/fgene.2016.00147](https://doi.org/10.3389/fgene.2016.00147)
35. Ivaska J, Pallari HM, Nevo J, Eriksson JE. Novel functions of vimentin in cell adhesion, migration, and signaling. *Exp Cell Res.* 2007;313(10):2050-2062. doi:[10.1016/j.yexcr.2007.03.040](https://doi.org/10.1016/j.yexcr.2007.03.040)
36. dos Santos G, Rogel MR, Baker MA, et al. Vimentin regulates activation of the NLRP3 inflammasome. *Nat Commun.* 2015;6:6574. doi:[10.1038/ncomms7574](https://doi.org/10.1038/ncomms7574)

CONFLICTS OF INTEREST

The authors have each completed and submitted an ICMJE form for disclosure of potential conflicts of interest. The authors declare that they have no competing interests, financial or otherwise, related to the current work. All the authors report that since the initial planning of the work, all support for the present manuscript was received from the Changzhou Key Laboratory of Maternal and Child Health Medicine and the Changzhou High-Level Health Talent Cultivation Project.

FUNDING

This research was supported by Changzhou Key Laboratory of Maternal and Child Health Medicine (CM20240013) and the Changzhou High-Level Health Talent Cultivation Project (2024CZBJ024).

ETHICAL APPROVAL AND INFORMED CONSENT

Ethical approval and informed consent were not required for this study as it is a secondary analysis of existing data.

DATA AVAILABILITY

The transcriptomic (GSE165004) and single-cell (GSE214607) datasets are publicly available in the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>). Summary statistics for Mendelian randomization were sourced from the IEU Open GWAS project (<https://gwas.mrcieu.ac.uk/>) and the FinnGen R12 release (https://www.finnngen.fi/en/access_results). Other web resources used include SwissTargetPrediction (<http://www.swisstargetprediction.ch/>), STITCH (<http://stitch.embl.de/>), DGIdb (<https://www.dgldb.org/>), CTD (<http://ctdbase.org/>), PDB (<https://www.rcsb.org/>), UniProt (<https://www.uniprot.org/>), PubChem (<https://pubchem.ncbi.nlm.nih.gov/>), and the CB-Dock2 server (<https://cadd.labshare.cn/cb-dock2/>). The data that support the findings of this study are available within the article or its [Supplementary file](#). Further inquiries can be directed to the corresponding author.

AUTHORS' CONTRIBUTIONS

KY and WZ: conception of the study. PZ and XS: analysis and manuscript preparation. HW and WZ: data analyses and writing of the manuscript. All authors read and approved the final version of the manuscript.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

SUPPLEMENTARY FILE DISCLAIMER

The content has been provided by the author(s) and has not been reviewed, verified, or endorsed by European Publishing. It may not have undergone peer review. The views, opinions, and recommendations expressed are solely those of the author(s) and do not necessarily reflect the position of European Publishing. European Publishing accepts no responsibility or liability for any consequences arising from the use of, or reliance on, this content.