ORIGINAL RESEARCH

Identifying and Validating GSTM5 as an Immunogenic Gene in Diabetic Foot Ulcer Using Bioinformatics and Machine Learning

Hongshuo Shi 🔟, Xin Yuan, Guobin Liu, Weijing Fan

Department of Peripheral Vascular Surgery, Institute of Surgery of Traditional Chinese Medicine, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, People's Republic of China

Correspondence: Weijing Fan; Guobin Liu, Email fan18811023202@126.com; 15800885533@163.com

Background: A diabetic foot ulcer (DFU) is a serious, long-term condition associated with a significant risk of disability and mortality. However, research on its biomarkers is still limited. This study utilizes bioinformatics and machine learning methods to identify immune-related biomarkers for DFU and validates them through external datasets and animal experiments.

Methods: This study used bioinformatics and machine learning to analyze microarray data from the Gene Expression Omnibus (GEO) database to identify key genes associated with DFU. Animal experiments were conducted to validate these findings. This research employs the datasets GSE68183 and GSE80178 retrieved from the GEO database as the training dataset for building a gene machine learning model, and after conducting differential analysis on the data, this study used package glmnet and package e1071 to construct LASSO and SVM-RFE machine learning models, respectively. Subsequently, we validated the model using the training set and validation set (GSE134431). We conducted enrichment analysis, including GSEA and GSVA, on the model genes. We also performed immune functional analysis and immune-related analysis on the model genes. Finally, we conducted immunohistochemistry (IHC) validation on the model genes.

Results: This study identifies GSTM5 as a potential immune-related key target in DFU using machine learning and bioinformatics methods. Subsequent validation through external datasets and IHC experiments also confirms GSTM5 as a critical biomarker for DFU. The gene may be associated with T cells regulatory (Tregs) and T cells follicular helper, and it influences the NF-κB, GnRH, and MAPK signaling pathway.

Conclusion: This study identified and validated GSTM5 as a biomarker for DFU. This finding may potentially provide a target for immune therapy for DFU.

Keywords: bioinformatics, machine learning, diabetic foot ulcer, LASSO, SVM-RFE, GSTM5

Introduction

Patients with newly diagnosed diabetes or a history of diabetes who develop peripheral arterial disease, peripheral neuropathy, infection, ulcers, neuro-osteoarthropathy, gangrene, or amputation of any kind in the foot are considered to have diabetes-related foot diseases. Diabetic foot ulcer (DFU) is one of the most serious complications. Patients with diabetic foot ulcers usually have one or more risk factors, among which diabetic peripheral neuropathy and peripheral arterial disease play important roles.[1] DFU represents one of the most catastrophic outcomes of diabetes and stands as a considerable healthcare challenge. It imposes serious economic, physical, and psychological burdens on patients.[2] Extensive alterations related to diabetes, such as neuropathy and vascular conditions, frequently worsen the advancement of DFU.[3] The lifetime occurrence of DFU among individuals with diabetes varies between 15% and 25%. Furthermore, in cases where there is a prior history of foot injuries or infections, the occurrence of DFU in diabetic patients rises to a range of 19% to 34%, with approximately 17% of these cases ultimately necessitating amputation.[4] In the year 2019, worldwide direct healthcare costs associated with diabetes amounted to around $700 billion. Projections suggest that by 2030, these expenses are anticipated to rise to $825 billion, with the medical costs linked to DFU constituting a third of

the overall expenditure for managing diabetes.[5] The development of DFU involves a notably intricate process, frequently triggered by distal neuropathy and peripheral vascular disorders in the lower limbs. In over half of DFU instances, infections are also present, adding complexity and difficulty to the treatment process.[6] Therefore, it is crucial to study new biomarkers of DFU and inhibit the development of chronic wounds in order to improve treatment efficiency and prognosis for DFU patients.[7]

The slow onset and delayed resolution of inflammatory response are among the major factors contributing to the difficulties in wound healing.[8] Immune cell infiltration is also a pivotal factor in the onset and progression of DFU. When viewed from a molecular standpoint, the process of wound healing takes place following a disruption in the skin's protective barrier. This process is typically facilitated by growth factors and cytokines released by specialized cells that become activated as part of the immune response.[9] Sawaya et al observed that in cases of DFU, the transcription factors FOMI1 and STAT3, which support the survival of immune cells, are suppressed. This ultimately hinders the process of wound healing in diabetic patients.[10] Through this study, our objective is to clarify the potential of immune-related biomarkers in enhancing the effectiveness of DFU treatment.

The high-throughput transcriptome sequencing data and clinical annotations provided by the DFU Initiative enable us to investigate the changes in transcriptional patterns and associated molecular pathways relevant to DFU in biological research.[11] Several research investigations have made use of gene expression data extracted from the Gene Expression Omnibus (GEO) to explore the molecular markers associated with the progression of DFU.[7,12] Due to the progress in high-throughput sequencing technologies and the integration of machine learning in the medical field, novel methodologies have surfaced for examining molecular targets in diverse diseases.[11,13] Through the application of bioinformatics techniques, we acquired a gene expression dataset for DFU from the GEO database. In order to pinpoint biomarkers linked to DFU, we employed two machine learning algorithms: LASSO and SVM-RFE. Following that, candidate genes strongly associated with immune infiltration were validated using a separate, independent validation dataset. The glutathione S-transferase (GST) gene family is one of the major xenobiotic detoxification enzymes that can protect cells from the harmful effects of toxic drugs and environmental electrophilic agents. GSTM5 has been found to be dysregulated in various tumors.[14] However, there is no relevant research indicating its role in DFU. Therefore, in this study, we constructed a mouse model of DFU and validated the GSTM5 gene. The flowchart of this study is shown in Figure 1.

Materials and Methods
Raw Data
Retrieval of microarray data on mRNA expression profiles related to DFU from the GEO database. This study utilized the GEO datasets GSE68183, GSE80178, and GSE134431, with their respective platform files GPL16686 and GPL18573. GSE68183 and GSE80178 are used as training datasets, while GSE134431 is used as a validation set. Detailed information about the aforementioned datasets can be found in Table 1.

Data Filtering and Processing
Transform the downloaded probe matrix into a gene expression matrix by referencing the probe annotation file. In cases where a gene corresponds to multiple probes, calculate the mean value of those probes to represent the ultimate expression level of the gene.[15] After standardizing datasets GSE68183 and GSE80178, the SVA package is used to address batch effects between GSE68183 and GSE80178. Principal component analysis (PCA) is employed to evaluate the removal of batch effects. Subsequently, the limma package is utilized to analyze the differential expression matrix comparing the control group and DFU. To identify genes related to immune infiltration in DFU patients, we employ the criteria of $|\log FC| > 1$ and an adjusted p-value < 0.05[16] as thresholds for detecting differentially expressed genes (DEGs).

Machine Learning
The LASSO method is used to identify essential features from large-scale data and screen gene models. The "glmnet" R package is utilized to implement this method in the process.[17] SVM-RFE is implemented through the "e1071"[18] R package and has been widely used as an effective feature selection algorithm for the selection of central diagnostic
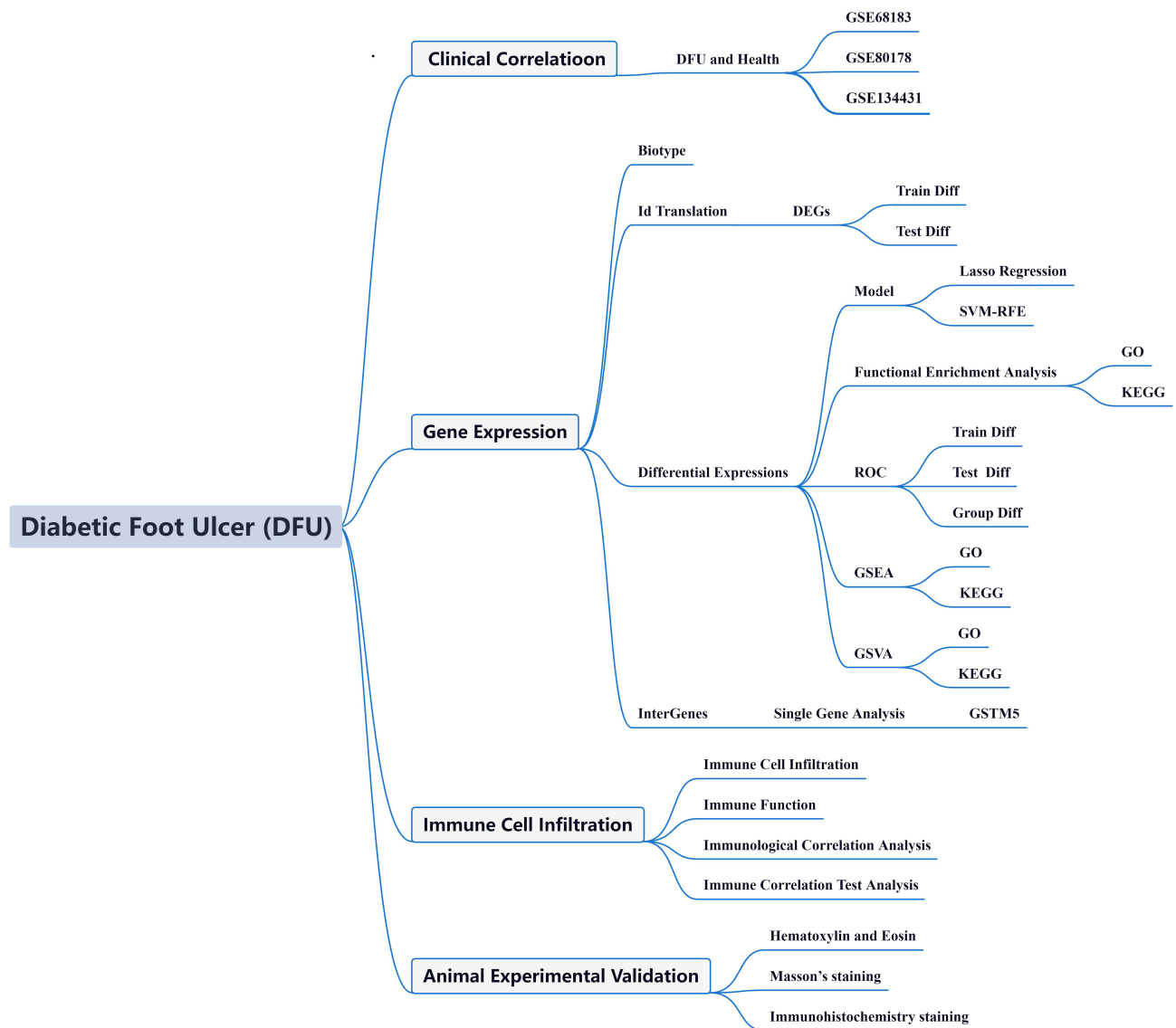
Figure 1 Research framework.

genes.[19] The genes derived from the aforementioned two machine learning models can be mapped and employed for subsequent analysis. The Receiver Operating Characteristic (ROC) curve evaluates the diagnostic utility of these biomarkers.

Pathway and Functional Enrichment Analysis

Apply Kyoto Encyclopedia of Genes and Genomes (KEGG)[20] and Gene Ontology (GO)[21] enrichment analysis using the R package clusterProfiler.[22] Additionally, we performed Gene Set Enrichment Analysis (GSEA) to identify potential

Table 1 Dataset Information

Dataset	Platform	Count	DFU	Control
GSE68183	GPL16686	6	3	3
GSE80178	GPL16686	12	9	3
GSE134431	GPL18573	21	13	8

pathways.[23] Gene Set Variation Analysis (GSVA) is an unsupervised, non-parametric analytical technique primarily employed to evaluate the enrichment of gene sets in sequencing data. GSVA enables the assessment of potential alterations in pathway activity within each individual sample.[24]

Immunocyte Infiltration Analysis

The relative abundance of 22 lymphocyte subtypes in each DFU and normal sample was calculated based on the CIBERSORT algorithm and the immune cell LM22 gene set.[25] The Wilcoxon test was used to compare the differences in 22 immune cell types between DFU and normal samples.

Correlation Analysis Between Diagnostic Biomarkers and Immune Cells

Performing Spearman correlation analysis to evaluate the correlation between diagnostic biomarkers and infiltrating immune cells.[26]

Establishment and Treatment of DFU Mouse Model

Male C57BL/6 mice of adult age (21–25 g, 8–10 weeks old) were procured from Shanghai Model Organisms Center Inc. These mice were accommodated at the Experimental Animal Center of Shanghai University of Traditional Chinese Medicine in an environment that met specific pathogen-free standards. They were maintained on a 12-hour light/dark cycle and provided unrestricted access to both water and food.[27] The mice were allocated randomly into distinct groups (n=4), which included the non-diabetic group treated with phosphate-buffered saline (PBS), and the DFU group. For a duration of three weeks, the DFU group of mice were fed a high-fat diet constituting 60% of their total calorie intake (FB-D12451, obtained from Wuxi Fan Bo Biotechnology Co., Ltd.), and within one week, they were subjected to intraperitoneal injections of STZ (40 mg/kg, Cat. No. 2196GR001, BioFRoxx).[28] Until the level of fasting blood glucose surpassed 16.7 mmol/L, it was classified as type 2 diabetes.[29] Following that, a full-thickness skin wound measuring 1*1cm2 in size, penetrating down to the fascial layer, was generated on the dorsal surface of the mouse by lifting the skin with forceps.[30] The wound was covered with a 1cm^2 gauze saturated with a wound dressing solution and replaced daily. To assess the condition of the dorsal wounds in mice, photographs of the wounds were taken on days 0, 7, and11 using a digital camera (Nikon, To to).[31] Quantification of wound closure rate was performed using ImageJ software (Bethesda, MD) and calculated as follows: t-t_0/t_0 × 100% (t: evaluation of wound healing, t0: initial injury).[27]

Hematoxylin and Eosin and Masson Staining

The wound tissue was preserved in 4% paraformaldehyde (PFA, catalog number P0099, Beyotime) for 48 hours. Following standard paraffin embedding and sectioning procedures, the wound tissue underwent staining with hematoxylin and eosin (H&E, catalog number C01105M, Beyotime, China) as well as Masson's trichrome stain (catalog number G1340, Solarbio Life Sciences, China). The stained tissues were subsequently examined utilizing a digital slide scanning system (point M8, Precipoint).

Immunohistochemistry Staining

The ulcerated tissues from both groups of mice were gathered and preserved in 4% paraformaldehyde for a duration of 24 hours. Subsequently, these tissues were encased in paraffin wax and later sliced into sections with a thickness of 5μm. Immunohistochemistry (IHC) analyses were conducted on these paraffin-embedded sections. The sections were subjected to an overnight incubation at 4°C with GSTM5 antibodies. Afterward, the slides designated for IHC staining underwent incubation with the secondary antibody (catalog number GB23303, Servicebio, diluted at 1:200). Visualization was achieved using DAB chromogen, followed by counterstaining with hematoxylin. Images were recorded using an optical microscope (Olympus BX41, Shanghai, China).

Statistical Analysis

Statistical assessment of disparities in gene expression between the two groups was conducted utilizing the ggpubr package within R software. All gene expression data and statistical analyses were executed using R version 4.3.1. If the

data conforms to a normal distribution, a two-sample independent *t*-test is employed to compare the two groups. In cases where the data does not conform to a normal distribution, the Wilcoxon test is utilized for between-group comparisons. A p-value below 0.05 is regarded as statistically significant.

Ethics Approval and Consent to Participate

All animal experiments were approved by the Review Board of Animal Ethics Committee of Shanghai University of Traditional Chinese Medicine (Approval No.: PZSHUTCM220711028). The study is reported in accordance with ARRIVE guidelines (https://arriveguidelines.org). The animal experimentation in this study was conducted in full compliance with the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978).

Results

GEO Data Processing

We have integrated two DFU datasets, GSE68183 and GSE80178, which collectively contain 6 normal skin samples and 12 DFU samples. As shown in Figure 2, the gene expression levels and PCA of each sample before (Figure 2A) and after (Figure 2B) eliminating batch effects.

Identifying DEGs Results in DFU

Differential analysis results indicate that there are a total of 234 DEGs in this study, with 59 upregulated and 175 downregulated. DEGs can be visualized through heatmaps (Figure 2C) and volcano plots (Figure 2D).

Screening Disease-Specific Genes

In this study, we employed LASSO and SVM-RFE models to identify DFU feature genes. In the LASSO model, the λ value of 7 was selected as the optimal variable selection criterion, resulting in the identification of genes LGR5, GSTM5, LCE3D, TRMT1L, XDH, and SRY (Figure 3A and B) as the feature genes. Based on the SVM-RFE algorithm, genes GSTM5, CBR4, KLHDC1, ZBED8, ENPP5, ZNF248, CCDC15, TRMT1L, MATN2, CSTA, TMEM45B, NRN1, and FLRT2 (Figure 3C and D) were identified as the feature genes. Subsequently, the above two gene models were mapped, resulting in the final identification of genes GSTM5 and TRMT1L (Figure 3E).

Analysis Disease-Specific Genes

In the training set, the violin plot reflects the expression levels of genes GSTM5 and TRMT1L in the DFU group and the control group. As shown in the figure, both genes GSTM5 (Figure 4A) and TRMT1L (Figure 4B) are down-regulated in the DFU group. The line graph illustrates the expression profiles of genes GSTM5 and TRMT1L across various samples (Figure 4C). Furthermore, the ROC curve demonstrates that both genes GSTM5 (Figure 4D) and TRMT1L (Figure 4E) have good diagnostic performance.

In the validation group, only GSTM5 (Figure 5A) showed significant down-regulation in the DFU group. However, there were no differences observed between the two groups in terms of the TRMT1L gene (Figure 5B). Furthermore, the ROC curve also indicates that the diagnostic performance of GSTM5 (Figure 5C) is superior to TRMT1L (Figure 5D).

Analysis of Gene Differential Expression and Enrichment Analysis in the Model

Based on the high and low expression of GSTM5, we conducted differential analysis of the genes in the training set. The results of the differential analysis were visualized as a heatmap (Figure 6A) and correlation heatmap (Figure 6B). Enrichment analysis indicates that the genetic biological processes of differentially expressed genes are concentrated in cell-substrate adhesion (GO:0031589), epithelial cell proliferation (GO:0050673), and negative regulation of leukocyte apoptotic process (GO:2000107) (Figure 6C). KEGG enrichment analysis indicates that the differentially expressed genes are primarily involved in NF-κB signaling pathway (hsa04064), Nitrogen metabolism (hsa00910), and GnRH signaling pathway (hsa04912) (Figure 6D). These results indicate that GSTM5 is closely associated with inflammation. GSEA
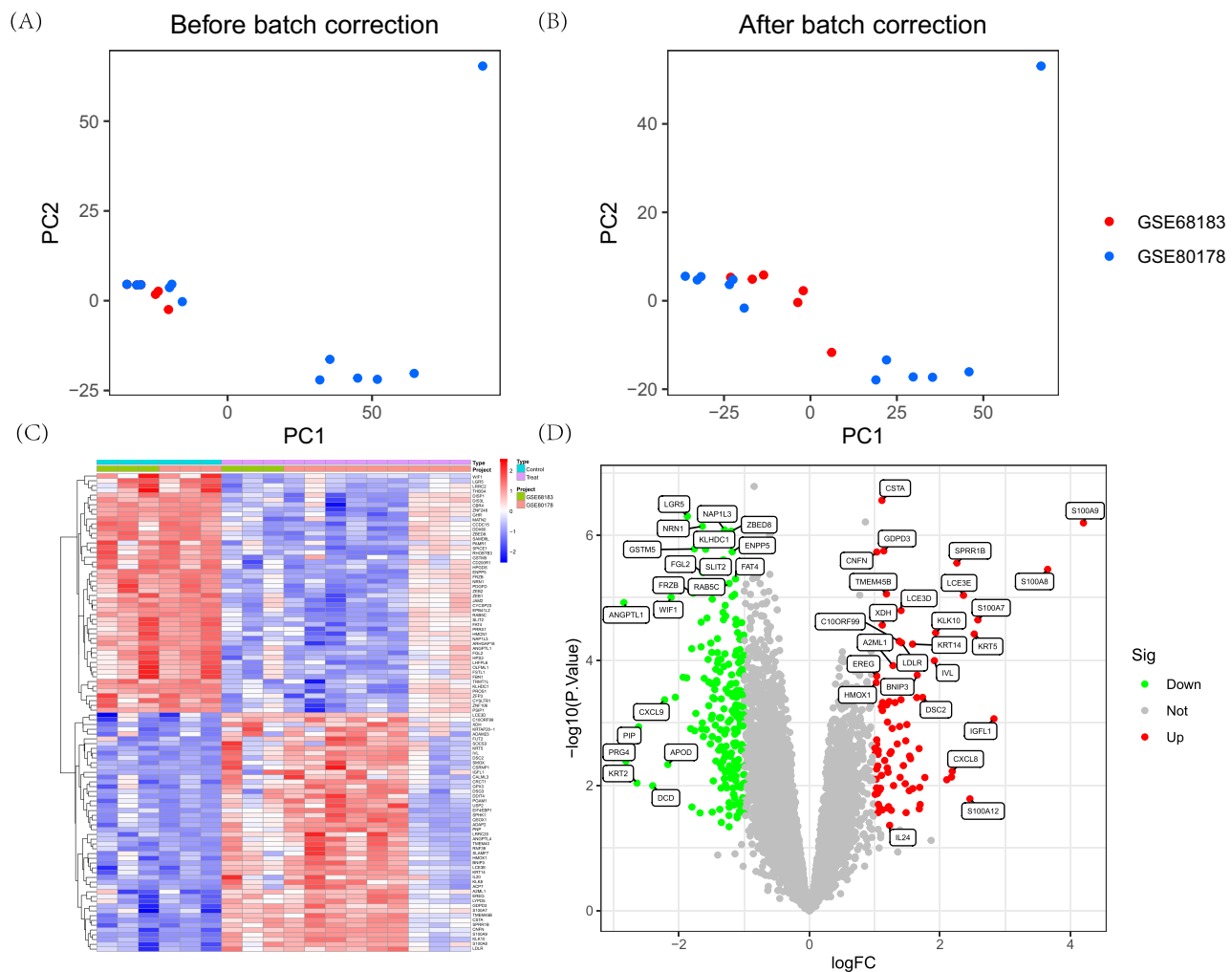
Figure 2 Dataset preprocessing and difference analysis. (**A**) Principal component analysis (PCA) analysis before batch effect removal; (**B**) PCA analysis after batch effect removal; (**C**) Heatmap of differential analysis between diabetic foot ulcer (DFU) group and control group; (**D**) Volcano plot of differential analysis between DFU group and control group.

analysis indicates that in the high-expression group of GSTM5, butanoate metabolism, complement and coagulation cascades, and drug metabolism cytochrome p450 are active (Figure 6E). In the low-expression group of GSTM5, insulin signaling pathway, MAPK signaling pathway, and neurotrophin signaling pathway are active (Figure 6F). GSVA analysis also indicates a close association between GSTM5 and terpenoid backbone biosynthesis and complement and coagulation cascades (Figure 6G).

Analysis of Immune Function and Immune Cell Infiltration

The immune function analysis suggests that GSTM5 may have a certain relationship with B cells and Cytolytic activity (Figure 7A). In DFU samples, B cells naive and Dendritic cells activated are higher compared to the control group, while T cells CD4 memory resting, T cells regulatory (Tregs), T cells follicular helper, Macrophages M1, and Mast cells resting are lower compared to the control group (Figure 7B). The distribution of immune cells and the PCA analysis are shown in the Figure 7C. The distribution of immune cells is shown in Figure 7D. Further investigation of the relationship between 22 immune cells in all samples revealed that Macrophages M1 and T cells CD4 memory resting show a strong positive correlation, while B cells naive and Macrophages M2 exhibit a strong negative correlation (Figure 7E). The correlation analysis between gene A and immune cells is shown in Figure 8A. We found that GSTM5 is significantly positively correlated with T cells regulatory (Tregs) (r = 0.66, p=0.019) (Figure 8B) and T cells follicular helper (r = 0.59, p=0.041) (Figure 8C).
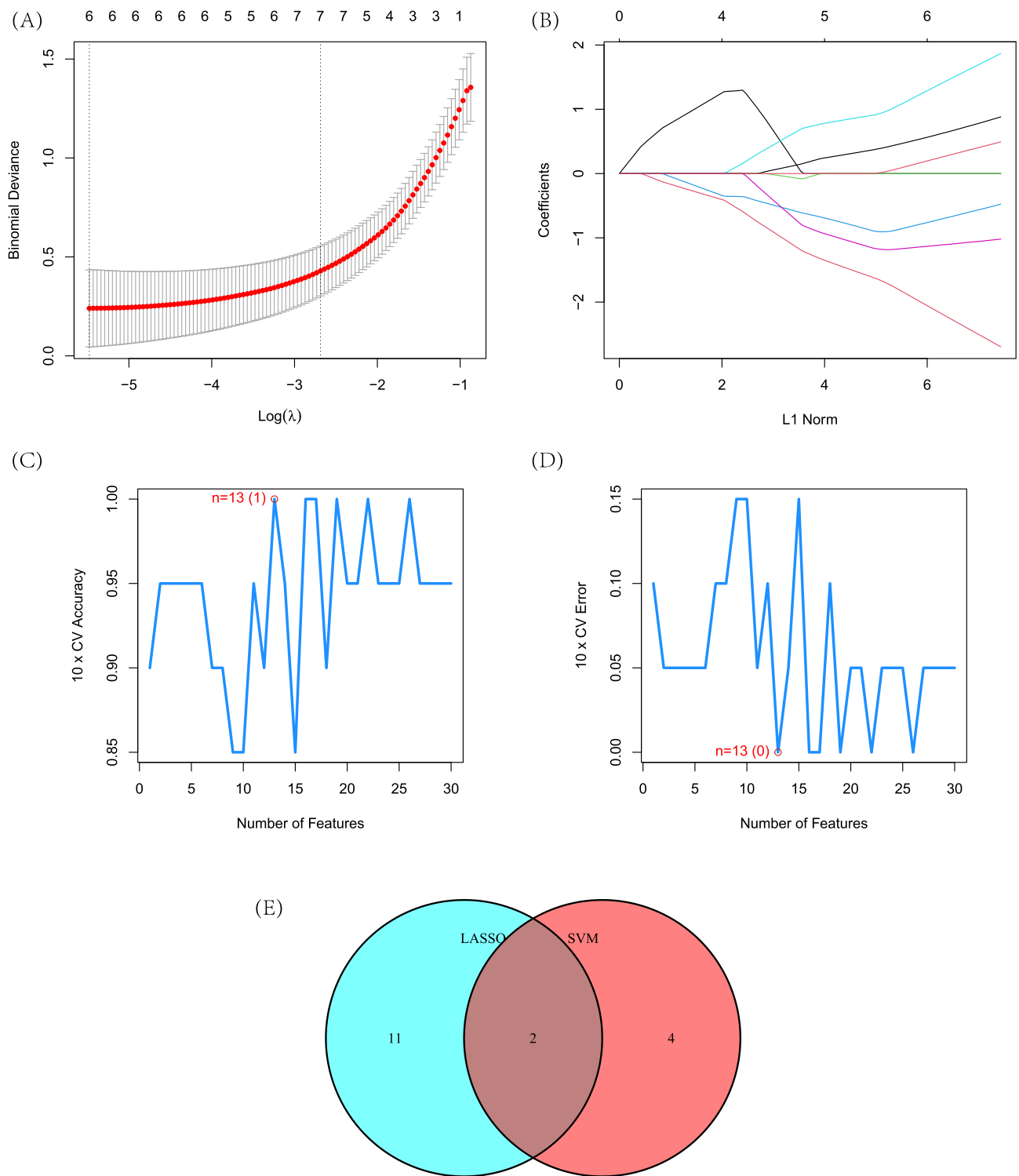
Figure 3 Construction of machine learning models. (**A**) Construction of Lasso model; (**B**) Variation of hub gene coefficients with different logarithmic λ values; (**C**) and (**D**) Construction of model SVM-RFE; (**E**) Mapping between Lasso and SVM-RFE models.

Gene Expression in DFU

The results revealed that in the control group, the scabs on the wound fell off, the wound size decreased, and there was redness. In the DFU group, the wound surface appeared yellow-brown scabs with a small amount of thin purulent discharge (Figure 9A), and the line graph of wound size measurement is found in Supplementary Material 1. Histological staining with
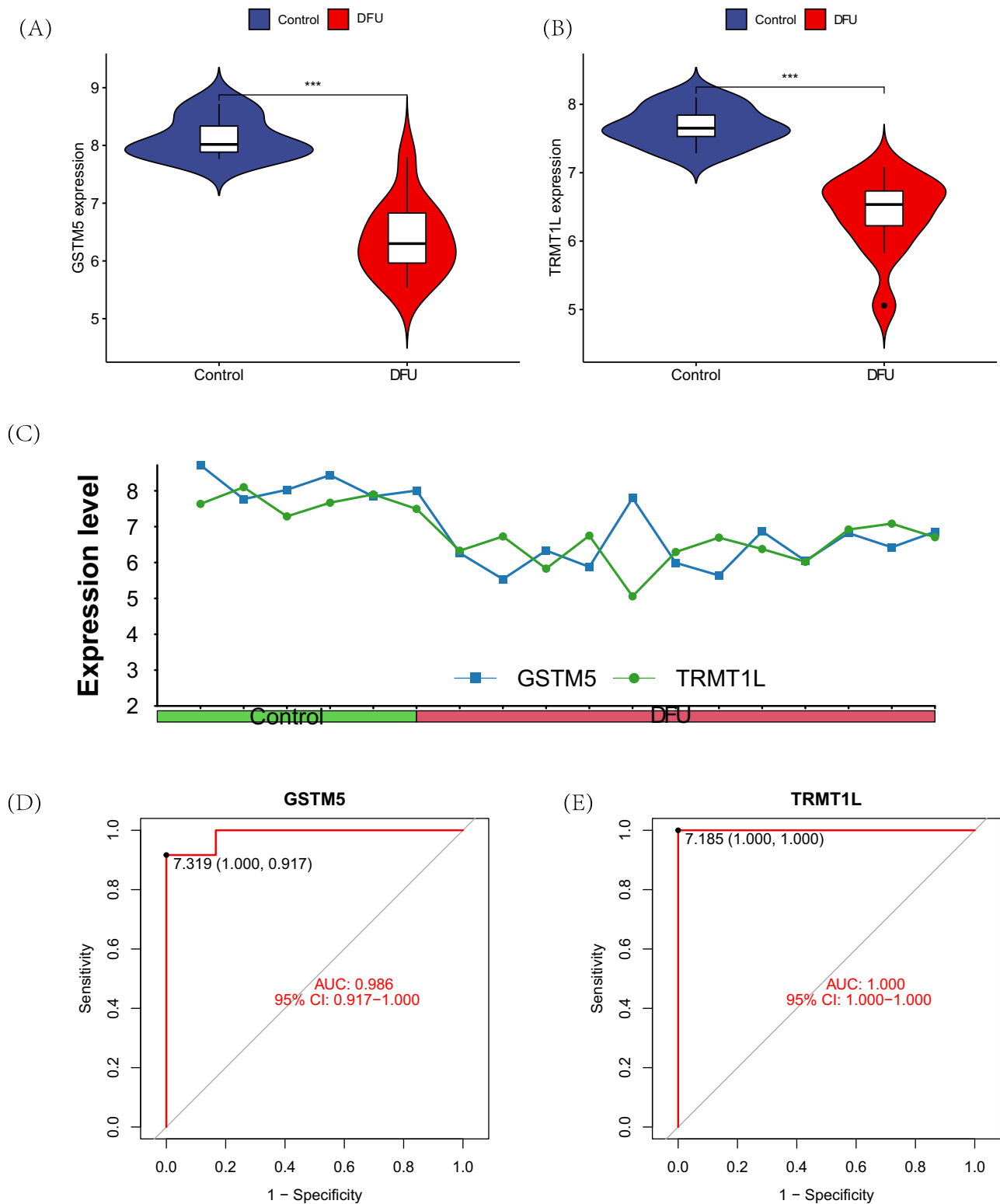
Figure 4 Differences in the intersection model within the training set. (**A**) The differential expression of gene GSTM5 between the DFU group and the control group; (**B**) The differential expression of gene TRMT1L between the DFU group and the control group; (**C**) The expression profiles of genes GSTM5 and TRMT1L in each sample; (**D**) The ROC curve of GSTM5; (**E**) The ROC curve of TRMT1L.
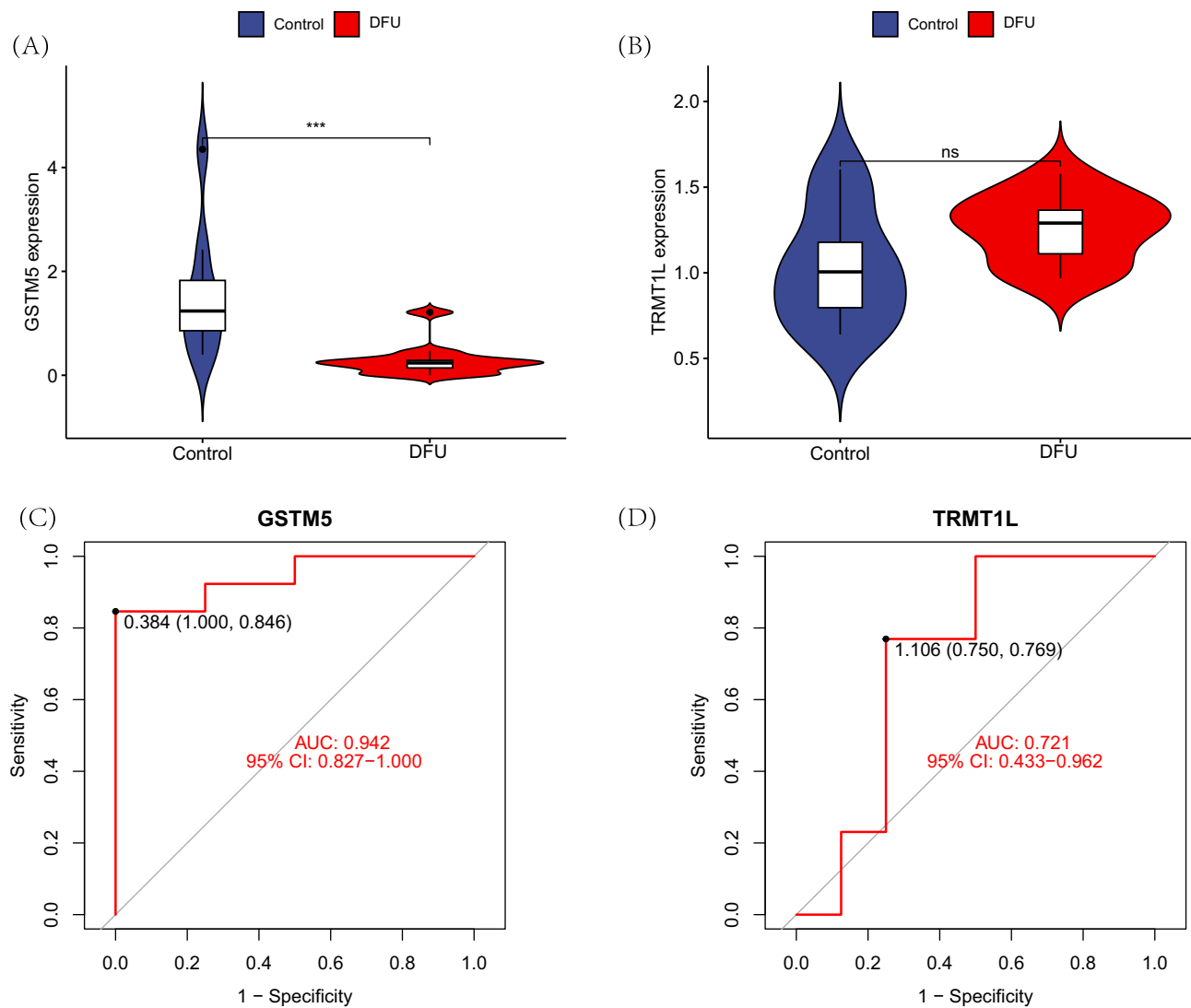**Note**: ***P<0.001.

6248

Journal of Inflammation Research 2023:16

Figure 5 Differences in the intersection model within the validation set. (**A**) The differential expression of gene GSTM5 between the DFU group and the control group; (**B**) The differential expression of gene TRMT1L between the DFU group and the control group; (**C**) The ROC curve of GSTM5; (**D**) The ROC curve of TRMT1L. **Note**: ***P<0.001.

HE (Hematoxylin and Eosin) and Masson's trichrome revealed that in the control group, there was epidermal regeneration with a small amount of fibroblasts and collagen, and a large number of capillaries in the dermis. In the DFU group, the skin appeared thinner with disordered and reduced epidermal cell arrangement, scattered fibroblasts, fewer newly formed blood vessels, reduced lumens, and extensive infiltration of inflammatory cells (Figure 9B). In conclusion, these results indicate the successful establishment of the DFU model. The IHC results demonstrated that the control group exhibited positive expression of GSTM5, while the DFU group showed significant inhibition (Figure 9C).

Discussion

DFU is one of the most common complications of diabetes, causing not only burdens for patients but also posing challenges for overall health, nursing practices, and the social environment.[32] Patients with diabetes often experience weakened immune systems, and when coupled with the vulnerability of the feet as one of the body's vital organs, infections can exacerbate inflammation and ultimately play a role in the onset of DFU.[33] This study, from an immunological perspective, identified GSTM5 as a potential key target for DFU using machine learning and
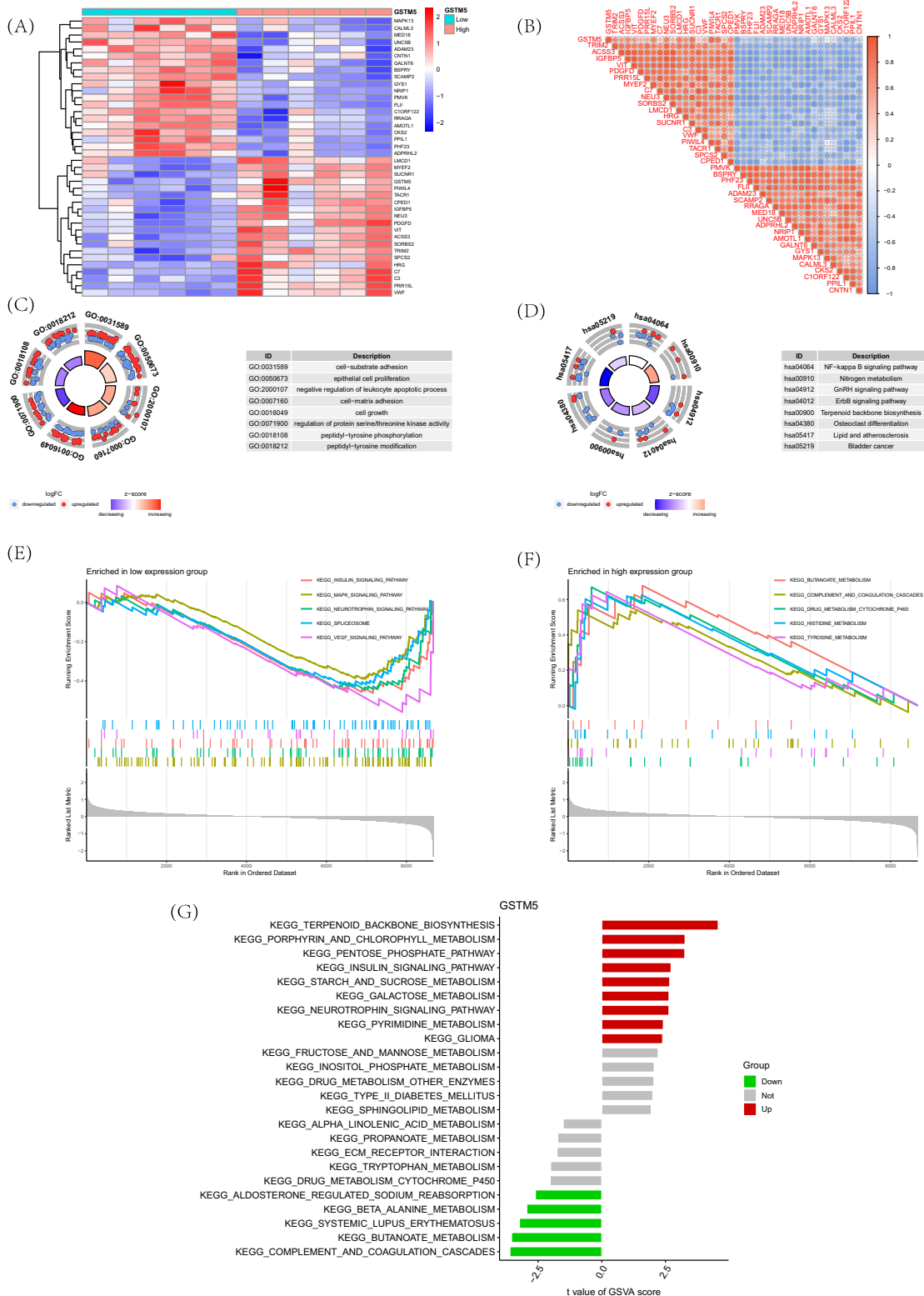
Figure 6 Differential analysis and enrichment analysis of genes related to GSTM5. (**A**) Differential analysis heatmap; (**B**) Correlation analysis heatmap; (**C**) Differential gene GO analysis; (**D**) Differential gene KEGG analysis; (**E**) GSEA analysis of downregulated genes; (**F**) GSEA analysis of upregulated genes; (**G**) GSTM5-related GSVA analysis.
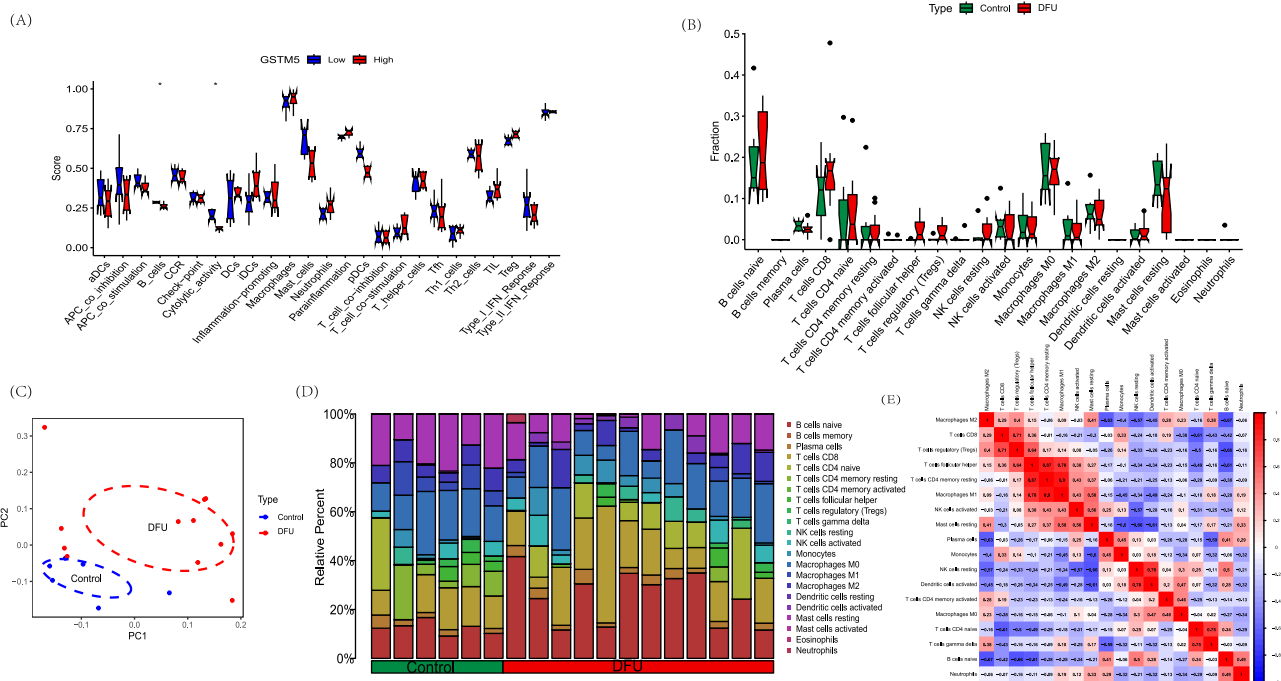
Figure 7 Immune analysis. (**A**) Immune functional analysis of the GSTM5 gene; (**B**) DFU immune infiltration analysis; (**C**) PCA analysis of immune infiltration results; (**D**) Immune cell distribution; (**E**) Immune cell correlation analysis.
Notes: *p<0.05.

bioinformatics methods. The low expression of GSTM5 in both the validation set and animal experiments among the DFU group suggests the reliability of the findings in this study.

Based on the conclusions of this study, GSTM5 is significantly underexpressed in DFU tissues, suggesting that it may be a potential protective factor in DFU patients. GSTM5 is a member of the Mu class within the glutathione S-transferase (GST) family, and these family members are distributed across a 97 kb region on chromosome 1p13.[34] The GST gene family is one of the major xenobiotic detoxification enzymes, protecting cells from the harmful effects of toxic drugs and environmental electrophiles.[35,36] Being a member of the glutathione S-transferase family, GSTM5 is linked to multiple types of cancers, such as breast cancer, prostate cancer, ovarian cancer, and colorectal cancer.[37] Earlier research has suggested that GSTM5 expression is reduced in bladder cancer tissues when compared to normal tissues. Furthermore, it has been demonstrated that elevated GSTM5 levels can lead to a reduction in the proliferation, migration, and colony-forming capacity of bladder cancer cells. These findings imply that GSTM5 may effectively serve as a tumor suppressor in bladder cancer, offering a substantial protective effect.[38] Furthermore, studies have also indicated that GSTM5 has low expression levels in breast cancer,[38] prostate cancer,[39] and Barrett's adenocarcinoma.[40] GSTM5 encodes a protein belonging to the glutathione S-transferase family, and these proteins are involved in limiting oxidative damage to tissues.[41] Oxidative damage is a key driving factor in the development of DFU.[42] In the current study, GSTM5 has shown a significant protective effect in most diseases, and our IHC testing also indicates its low expression in DFU tissues.

The inflammatory response, encompassing early reactions following an injury,[43] holds a pivotal role in facilitating the healing process. The recruitment of immune cells to the injury site aids in orchestrating multifaceted healing responses among cells and serves to thwart infections.[44,45] The process of acute wound healing comprises several sequential and overlapping stages, including hemostasis, inflammation, proliferation, and remodeling.[43] The inflammatory response is often regarded as the "driving force" that initiates the repair process, and because of its capacity for potential harm, it is closely controlled and regulated. The results of immune infiltration indicate that in the DFU group, B cells naive and dendritic cells are highly expressed, while in the control group, T cells CD4 memory resting, T cells regulatory (Tregs),
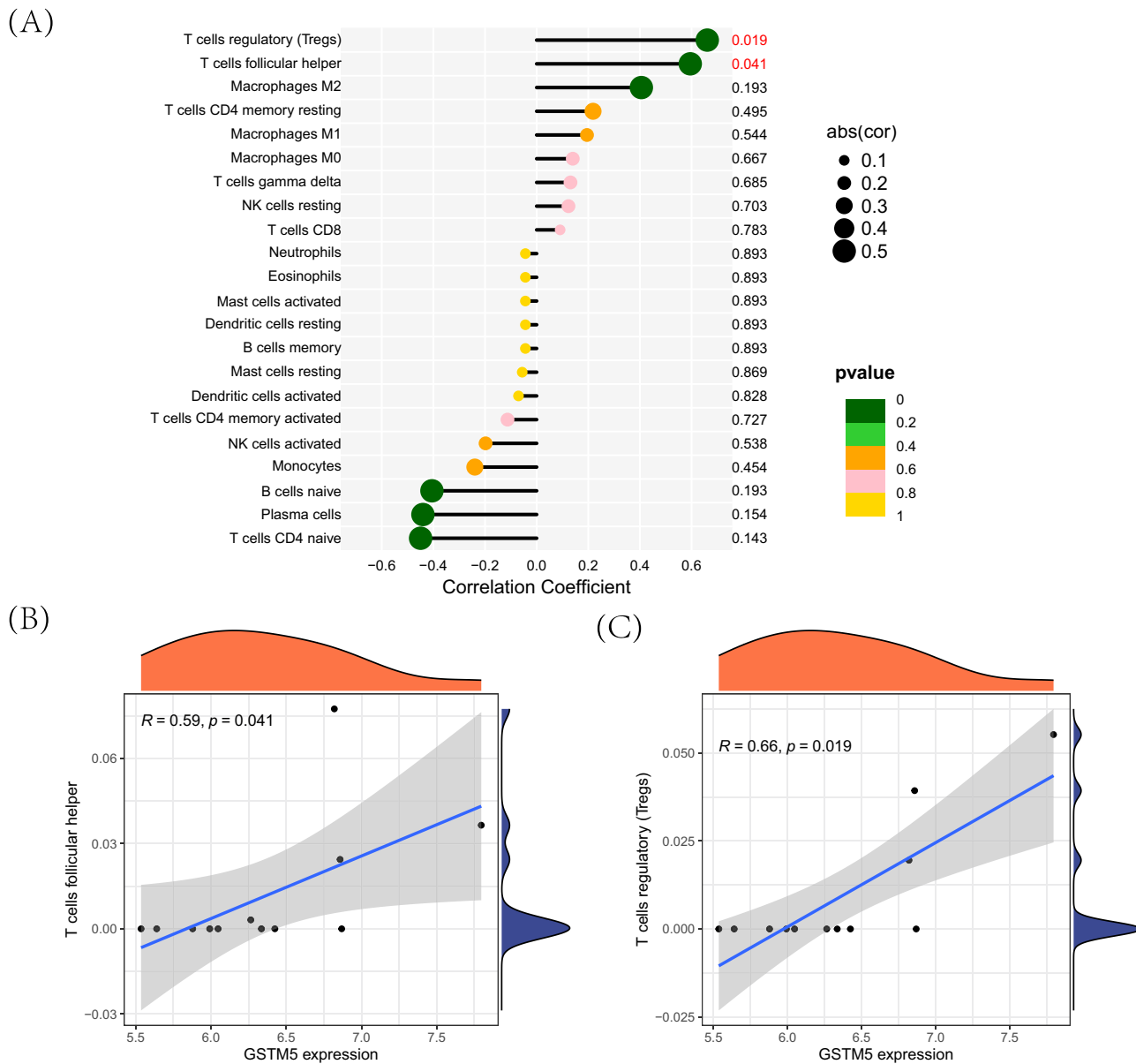
(A)

(B) (C)

Figure 8 Immune cell correlation analysis of GSTM5. (**A**) Correlation analysis; (**B**) GSTM5 correlation analysis with T cells follicular helper; (**C**) GSTM5 correlation analysis with T cells regulatory (Tregs).

T cells follicular helper, Macrophages M1, and Mast cells resting are highly expressed. The ssGSEA results also suggest that GSTM5 regulates immune functions related to B cells and cytolytic activity, and the immune cells T cells regulatory (Tregs) and T cells follicular helper were found to be positively correlated with GSTM5 in this study. Studies suggest that the timing and effectiveness of wound healing can be deliberately manipulated through the direct application of mature, unactivated B cells. This strategy can successfully modify the equilibrium of mature immune cell populations within the wound microenvironment, thereby expediting the healing process.[46] Tregs, a distinctive and vital subset of T cells, are identified by their expression of Foxp3 and are classified as CD4+CD25+ cells. They play a unique and essential role in immune surveillance.[47,48] Studies have indicated that T cells migrate to the site of injury during the proliferative phase.[49] According to studies, dendritic epidermal T cells have been demonstrated to enhance impaired wound healing in diabetic mice. They achieve this by promoting the migration and proliferation of keratinocytes while reducing keratinocyte
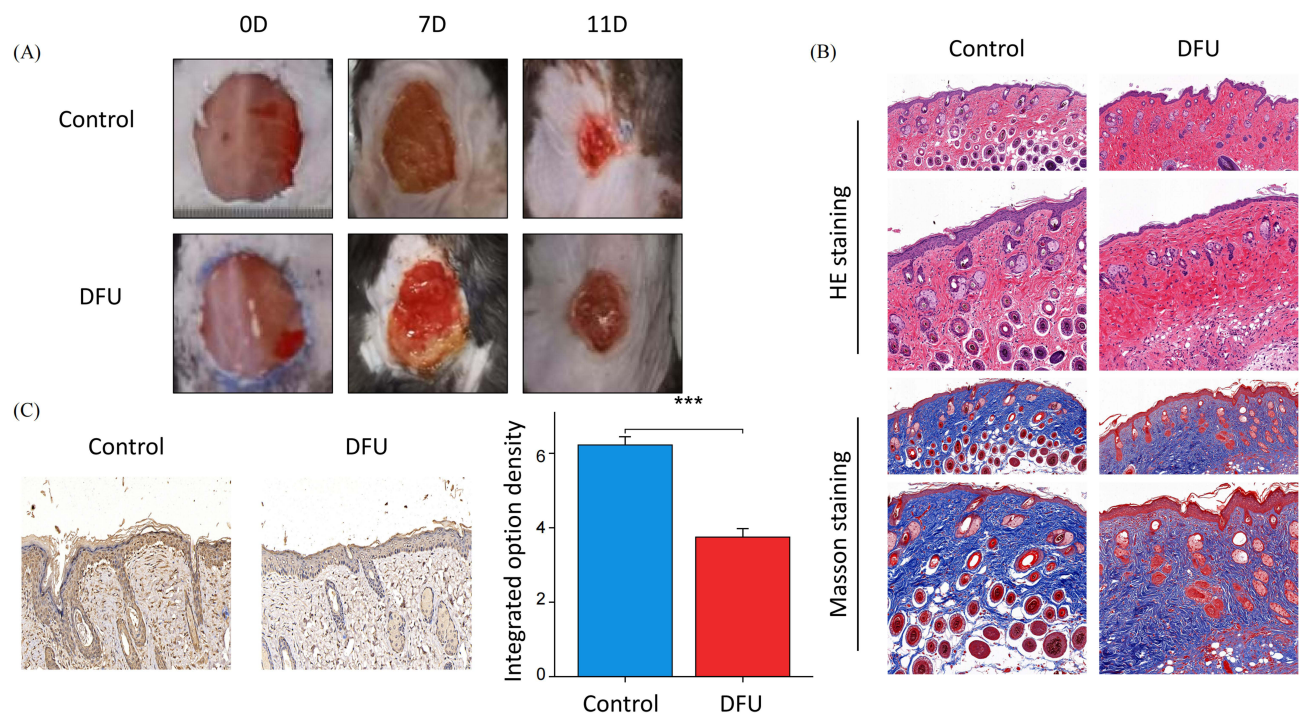
Figure 9 Animal experimental validation. (**A**) Optical pictures of the wound closure rate in the Control group and DFU group at days 0, 7, and 11 after the skin operation; (**B**) H&E and Masson staining images of wound tissues in the Control group and DFU group at day 7; (**C**) Representative images of immunohistochemical staining of GSTM5. **Note**: ***P<0.001.

apoptosis in an environment resembling diabetes.[50] Additionally, studies have indicated that factors secreted by T cells may help drive macrophage polarization during the wound healing process.[51]

The enrichment analysis suggests that GSTM5 may be correlated with NF-κB, GnRH, and MAPK signaling pathway. Elevated expression of NF-κB genes or proteins can result in increased oxidative stress, potentially impacting the wound healing process in diabetic patients.[8,52] Current understanding suggests that in diabetic wounds, the persistent elevation of TNF-α triggers NF-κB activation, which in turn causes the release of numerous inflammatory cytokines, thereby sustaining an ongoing inflammatory response. Consequently, various chemotactic factors and cytokines essential for wound healing lose their effectiveness, resulting in prolonged healing durations.[53] As per available information, the use of GnRH agonists is linked to an increased risk of diabetes.[54] Research suggests that in diabetic models, mice treated with GnRH agonists exhibit more delayed wound healing compared to untreated control mice, and treatment with GnRH antagonists inhibits the damage caused by GnRH agonist treatment.[55] The activation of the MAPK signaling pathway acts as a critical junction for numerous trauma-induced signaling pathways, facilitating wound healing. Nevertheless, during the wound healing process in diabetes, the local hyperglycemic condition can disrupt normal repair mechanisms. Adequate keratinocyte migration is essential for wound healing, yet high glucose levels hinder keratinocyte migration by suppressing the p38 MAPK/autophagy pathway. This interference results in challenges in wound healing.[56] Studies have indicated that the hyperglycemic condition in individuals with diabetes activates the MAPK pathway through both protein kinase and non-protein kinase pathways. An upregulation of MAPK is associated with the progression of DFU. Inhibiting the MAPK pathway can offer therapeutic benefits.[57]

Indeed, there is still a relative lack of research on biomarkers in the context of DFU. Recently, the combination of bioinformatics and machine learning analysis has emerged as a valuable tool for discovering novel biomarkers in DFU.[58] A previous study suggested that MAPK3 may be a biomarker associated with iron death in DFU.[7] It is worth noting that there is currently a lack of research in the literature concerning the immune system and its relevance to DFU. This study combines machine learning and immune infiltration methods to re-identify GSTM5 while also providing insights into its related functions and mechanisms. Additionally, the low expression of GSTM5 in both the validation set and animal experiments among the DFU group suggests the reliability of the findings in this study. While encouraging results have

been found, it is important to acknowledge the limitations of this study. To begin with, the clinical data utilized in this study were sourced from public databases, and the available clinical information for the samples was incomplete. This included limitations such as the unavailability of clinical pathological characteristics for the GSE series. Secondly, the data employed in this study relied on RNA sequencing. While animal experiments corroborated the findings of the bioinformatics analysis, it is imperative to validate the reproducibility and generalizability of these results in future investigations using clinical samples. However, this was not feasible within the constraints of the limited timeframe, as clinical samples were not readily accessible. Lastly, even if diabetic rats were created, the wound model created in these rats is an acute wound model, not a chronic wound.

Conclusion

In conclusion, this study identifies GSTM5 as a potential immune-related key target in DFU using machine learning and bioinformatics methods. Subsequent validation through external datasets and animal experiments also confirms GSTM5 as a critical biomarker for DFU.

Data Sharing Statement

All generated data have been included in the published manuscript. The datasets of gene expression generated during the current study are available in the NCBI Gene Expression Omnibus (GEO) database (accession number: GSE68183, GSE80178, and GSE134431). For all other information, contact with the author (Hongshuo Shi, jf17510413109@163.com).

Funding

Disclosure

The authors declare no competing interests.

References

1. Schaper NC, van Netten JJ, Apelqvist J, et al. Practical guidelines on the prevention and management of diabetes-related foot disease (IWGDF 2023 update). *Diabetes Metab Res Rev*. 2023. doi:10.1002/dmrr.3657
2. Zou J, Zhang W, Chen X, Su W, Yu D. Data mining reveal the association between diabetic foot ulcer and peripheral artery disease. *Front Public Health*. 2022;10:963426. doi:10.3389/fpubh.2022.963426
3. Xiong Y, Chen L, Yan C, et al. Circulating exosomal miR-20b-5p inhibition restores Wnt9b signaling and reverses diabetes-associated impaired wound healing. *Small*. 2020;16(3):e1904044. doi:10.1002/smll.201904044
4. Du Y, Wang J, Fan W, Huang R, Wang H, Liu G. Preclinical study of diabetic foot ulcers: from pathogenesis to vivo/vitro models and clinical therapeutic transformation [published online ahead of print, 2023 Jul 12]. *Int Wound J*. 2023. doi:10.1111/iwj.14311
5. Williams R, Karuranga S, Malanda B, et al. Global and regional estimates and projections of diabetes-related health expenditure: results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2020;162:108072. doi:10.1016/j.diabres.2020.108072
6. Prompers L, Huijberts M, Apelqvist J, et al. High prevalence of ischaemia, infection and serious comorbidity in patients with diabetic foot disease in Europe. Baseline results from the Eurodiale study. *Diabetologia*. 2007;50(1):18–25. doi:10.1007/s00125-006-0491-1
7. Wang X, Jiang G, Zong J, et al. Revealing the novel ferroptosis-related therapeutic targets for diabetic foot ulcer based on the machine learning. *Front Genet*. 2022;13:944425. doi:10.3389/fgene.2022.944425
8. Zhou X, Guo Y, Yang K, Liu P, Wang J. The signaling pathways of traditional Chinese medicine in promoting diabetic wound healing. *J Ethnopharmacol*. 2022;282:114662. doi:10.1016/j.jep.2021.114662
9. Miao F, Li X, Wang C, Yuan H, Wu Z. Bioinformatics analysis of differentially expressed genes in diabetic foot ulcer and preliminary experimental verification. *Ann Transl Med*. 2023;11(2):89. doi:10.21037/atm-22-6437
10. Sawaya AP, Stone RC, Brooks SR, et al. Deregulated immune cell recruitment orchestrated by FOXM1 impairs human diabetic wound healing. *Nat Commun*. 2020;11(1):4678. doi:10.1038/s41467-020-18276-0

11. Wu Z, Liu P, Huang B, et al. A novel Alzheimer's disease prognostic signature: identification and analysis of glutamine metabolism genes in immunogenicity and immunotherapy efficacy. *Sci Rep*. 2023;13(1):6895. doi:10.1038/s41598-023-33277-x

12. Li Y, Ju S, Li X, et al. Characterization of the microenvironment of diabetic foot ulcers and potential drug identification based on scRNA-seq. *Front Endocrinol (Lausanne)*. 2023;13:997880. doi:10.3389/fendo.2022.997880

13. Wu Z, Li X, Gu Z, Xia X, Yang J. Pyrimidine metabolism regulator-mediated molecular subtypes display tumor microenvironmental hallmarks and assist precision treatment in bladder cancer. *Front Oncol*. 2023;13:1102518. doi:10.3389/fonc.2023.1102518

14. Hao X, Zhang J, Chen G, Cao W, Chen H, Chen S. Aberrant expression of GSTM5 in lung adenocarcinoma is associated with DNA hypermethylation and poor prognosis. *BMC Cancer*. 2022;22(1):685. doi:10.1186/s12885-022-09711-0

15. Bao W, Wang L, Liu X, Li M. Predicting diagnostic biomarkers associated with immune infiltration in Crohn's disease based on machine learning and bioinformatics. *Eur J Med Res*. 2023;28(1):255. doi:10.1186/s40001-023-01200-9

16. Zhu E, Shu X, Xu Z, et al. Screening of immune-related secretory proteins linking chronic kidney disease with calcific aortic valve disease based on comprehensive bioinformatics analysis and machine learning. *J Transl Med*. 2023;21(1):359. doi:10.1186/s12967-023-04171-x

17. Gao J, Kwan PW, Shi D. Sparse kernel learning with LASSO and Bayesian inference algorithm. *Neural Netw*. 2010;23(2):257–264. doi:10.1016/j.neunet.2009.07.001

18. Yoon S, Kim S. AdaBoost-based multiple SVM-RFE for classification of mammograms in DDSM. *BMC Med Inform Decis Mak*. 2009;9 Suppl 1 (Suppl 1):S1. doi:10.1186/1472-6947-9-S1-S1

19. Lin X, Yang F, Zhou L, et al. A support vector machine-recursive feature elimination feature selection method based on artificial contrast variables and mutual information. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2012;910:149–155. doi:10.1016/j.jchromb.2012.05.020

20. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–D462. doi:10.1093/nar/gkv1070

21. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. 2019;47(D1):D330–D338. doi:10.1093/nar/gky1055

22. 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

23. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–15550. doi:10.1073/pnas.0506580102

24. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;14:7. doi:10.1186/1471-2105-14-7

25. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–259. doi:10.1007/978-1-4939-7493-1_12

26. Zhou J, Huang J, Li Z, et al. Identification of aging-related biomarkers and immune infiltration characteristics in osteoarthritis based on bioinformatics analysis and machine learning. *Front Immunol*. 2023;14:1168780. doi:10.3389/fimmu.2023.1168780

27. Wang J, Wang Y, Huang R, et al. Uncovering the pharmacological mechanisms of Zizhu ointment against diabetic ulcer by integrating network analysis and experimental evaluation in vivo and in vitro. *Front Pharmacol*. 2022;13:1027677. doi:10.3389/fphar.2022.1027677

28. Cheng Y, Peng L, Deng X, et al. Prostaglandin F2α protects against pericyte apoptosis by inhibiting the PI3K/Akt/GSK3β/β-catenin signaling pathway. *Ann Transl Med*. 2021;9(12):1021. doi:10.21037/atm-21-2717

29. Du Y, Chen W, Li Y, Liang D, Liu G. Study on the regulatory effect of Panax notoginseng saponins combined with bone mesenchymal stem cell transplantation on IRAK1/TRAF6-NF-κB pathway in patients with diabetic cutaneous ulcers. *J Orthop Surg Res*. 2023;18(1):80. doi:10.1186/s13018-022-03467-w

30. Qiu Y, Wang Q, Chen Y, Xia S, Huang W, Wei Q. A novel multilayer composite membrane for wound healing in mice skin defect model. *Polymers (Basel)*. 2020;12(3):573. doi:10.3390/polym12030573

31. Huang S, Hu Z, Wang P, et al. Rat epidermal stem cells promote the angiogenesis of full-thickness wounds. *Stem Cell Res Ther*. 2020;11(1):344. doi:10.1186/s13287-020-01844-y

32. Subrata SA, Phuphaibul R. A nursing metaparadigm perspective of diabetic foot ulcer care. *Br J Nurs*. 2019;28(6):S38–S50. doi:10.12968/bjon.2019.28.6.S38

33. Chung HJ, Chun DI, Kang EM, et al. Trend and seasonality of diabetic foot amputation in South Korea: a population-based nationwide study. *Int J Environ Res Public Health*. 2022;19(7):4111. doi:10.3390/ijerph19074111

34. Lizard-Nacol S, Coudert B, Colosetti P, Riedinger JM, Fargeot P, Brunet-Lecomte P. Glutathione S-transferase M1 null genotype: lack of association with tumour characteristics and survival in advanced breast cancer. *Breast Cancer Res*. 1999;1(1):81–87. doi:10.1186/bcr17

35. Kim Y, Cha SJ, Choi HJ, Kim K. Omega class glutathione S-transferase: antioxidant enzyme in pathogenesis of neurodegenerative diseases. *Oxid Med Cell Longev*. 2017;2017:5049532. doi:10.1155/2017/5049532

36. Awasthi YC, Ramana KV, Chaudhary P, Srivastava SK, Awasthi S. Regulatory roles of glutathione-S-transferases and 4-hydroxynonenal in stress-mediated signaling and toxicity. *Free Radic Biol Med*. 2017;111:235–243. doi:10.1016/j.freeradbiomed.2016.10.493

37. Liu F, Xiao XL, Liu YJ, et al. CircRNA_0084927 promotes colorectal cancer progression by regulating miRNA-20b-3p/glutathione S-transferase mu 5 axis. *World J Gastroenterol*. 2021;27(36):6064–6078. doi:10.3748/wjg.v27.i36.6064

38. Jou YC, Wang SC, Dia YC, et al. Anti-cancer effects and tumor marker role of glutathione S-Transferase Mu 5 in human bladder cancer. *Int J Mol Sci*. 2021;22(6):3056. doi:10.3390/ijms22063056

39. Peng DF, Razvi M, Chen H, et al. DNA hypermethylation regulates the expression of members of the Mu-class glutathione S-transferases and glutathione peroxidases in Barrett's adenocarcinoma. *Gut*. 2009;58(1):5–15. doi:10.1136/gut.2007.146290

40. Sun C, Gu Y, Chen G, Du Y. Bioinformatics analysis of stromal molecular signatures associated with breast and prostate cancer. *J Comput Biol*. 2019;26(10):1130–1139. doi:10.1089/cmb.2019.0045

41. Strange RC, Spiteri MA, Ramachandran S, Fryer AA. Glutathione-S-transferase family of enzymes. *Mutat Res*. 2001;482(1–2):21–26. doi:10.1016/s0027-5107(01)00206-8

42. Lin S, Zhang Q, Li S, et al. Antioxidative and angiogenesis-promoting effects of tetrahedral framework nucleic acids in diabetic wound healing with activation of the Akt/Nrf2/HO-1 pathway. *ACS Appl Mater Interfaces*. 2020;12(10):11397–11408. doi:10.1021/acsami.0c00874

43. Eming SA, Martin P, Tomic-Canic M. Wound repair and regeneration: mechanisms, signaling, and translation. *Sci Transl Med*. 2014;6(265):265sr6. doi:10.1126/scitranslmed.3009337

44. Pastar I, Stojadinovic O, Yin NC, et al. Epithelialization in wound healing: a comprehensive review. *Adv Wound Care (New Rochelle)*. 2014;3 (7):445–464. doi:10.1089/wound.2013.0473

45. Gurtner GC, Werner S, Barrandon Y, Longaker MT. Wound repair and regeneration. *Nature*. 2008;453(7193):314–321. doi:10.1038/nature07039

46. Sîrbulescu RF, Boehm CK, Soon E, et al. Mature B cells accelerate wound healing after acute and chronic diabetic skin lesions. *Wound Repair Regen*. 2017;25(5):774–791. doi:10.1111/wrr.12584

47. Baekkevold ES, Wurbel MA, Kivisäkk P, et al. A role for CCR4 in development of mature circulating cutaneous T helper memory cell populations. *J Exp Med*. 2005;201(7):1045–1051. doi:10.1084/jem.20041059

48. Wang X, Fujita M, Prado R, et al. Visualizing CD4 T-cell migration into inflamed skin and its inhibition by CCR4/CCR10 blockades using in vivo imaging model. *Br J Dermatol*. 2010;162(3):487–496. doi:10.1111/j.1365-2133.2009.09552.x

49. Guo S, Dipietro LA. Factors affecting wound healing. *J Dent Res*. 2010;89(3):219–229. doi:10.1177/0022034509359125

50. Liu Z, Xu Y, Chen L, et al. Dendritic epidermal T cells facilitate wound healing in diabetic mice. *Am J Transl Res*. 2016;8(5):2375–2384.

51. Wolf SJ, Melvin WJ, Gallagher K. Macrophage-mediated inflammation in diabetic wound repair. *Semin Cell Dev Biol*. 2021;119:111–118. doi:10.1016/j.semcdb.2021.06.013

52. Deng L, Du C, Song P, et al. The role of oxidative stress and antioxidants in diabetic wound healing. *Oxid Med Cell Longev*. 2021;2021:8852759. doi:10.1155/2021/8852759

53. Zhang Q, Lenardo MJ, Baltimore D. 30 years of NF-κB: a blossoming of relevance to human pathobiology. *Cell*. 2017;168(1–2):37–57. doi:10.1016/j.cell.2016.12.012

54. Keating NL, O'Malley AJ, Smith MR. Diabetes and cardiovascular disease during androgen deprivation therapy for prostate cancer. *J Clin Oncol*. 2006;24(27):4448–4456. doi:10.1200/JCO.2006.06.2497

55. Lee YS, Kang SU, Lee MH, et al. GnRH impairs diabetic wound healing through enhanced NETosis. *Cell Mol Immunol*. 2020;17(8):856–864. doi:10.1038/s41423-019-0252-y

56. Papa S, Choy PM, Bubici C. The ERK and JNK pathways in the regulation of metabolic reprogramming. *Oncogene*. 2019;38(13):2223–2240. doi:10.1038/s41388-018-0582-8

57. Liu Y, Li Z, Li W et al. Discovery of β-sitosterol's effects on molecular changes in rat diabetic wounds and its impact on angiogenesis and macrophages. *Int Immunopharmacol*. 2023;126:111283. doi:10.1016/j.intimp.2023.111283

58. Wang L, Deng C, Wu Z, Zhu K, Yang Z. Bioinformatics and machine learning were used to validate glutamine metabolism-related genes and immunotherapy in osteoporosis patients. *J Orthop Surg Res*. 2023;18(1):685. doi:10.1186/s13018-023-04152-2