



## OPEN Integrative genetics and multiomics analysis reveal mechanisms and therapeutic targets in vitiligo highlighting JAK/STAT pathway regulation of CTSS

Zi-yue Dong<sup>1,2,5</sup>, Ming-jie He<sup>2,5</sup>, Yong-kai Yu<sup>3,5</sup>, Fang Wang<sup>2</sup>, Peng-yuan Zhao<sup>2</sup>, De-long Ran<sup>2</sup>, De-shuang Fu<sup>2</sup>, Qing He<sup>2</sup>, Run-ping Yang<sup>4</sup>✉ & Jiang-an Zhang<sup>2</sup>✉

Vitiligo is a complex autoimmune disease characterized by the loss of melanocytes, leading to skin depigmentation. Despite advances in understanding its genetic and molecular basis, the precise mechanisms driving vitiligo remain elusive. Integrating multiple layers of omics data can provide a comprehensive view of disease pathogenesis and identify potential therapeutic targets. The study aims to delineate the genetic and molecular mechanisms of vitiligo pathogenesis using an integrative multiomics strategy. We focus on exploring the regulatory influence of the JAK/STAT pathway on Cathepsin S, a potential therapeutic target in vitiligo. Our GWAS-meta analysis pinpointed five druggable genes: ERBB3, RHOH, CDK10, MC1R, and NDUFAF3, and underwent drug target exploration and molecular docking. SMR analysis linked CTSS, CTSH, STX8, KIR2DL3, and GRHRP to vitiligo through pQTL and eQTL associations. Microarray and single-cell RNA-seq data showed differential expression of CTSS and STAT1/3 in vitiligo patients' blood and skin lesions. Our study offers novel perspectives on vitiligo's genetic and molecular basis, highlighting the JAK/STAT pathway's role in regulating CTSS for antigen processing in melanocytes. Further research is needed to confirm these results and assess the therapeutic potential of CTSS and related genes.

**Keywords** Bioinformatics, CTSS, GWAS, Multiomics, Therapeutic targets, Vitiligo

Vitiligo is a prevalent autoimmune disorder characterized by the selective destruction of melanocytes, resulting in depigmented patches on the skin. In approximately 0.5–1% of the global population, vitiligo significantly impacts the quality of life and psychosocial well-being of affected individuals<sup>1,2</sup>. Despite extensive research, the etiology and pathogenesis of vitiligo remain incompletely understood, and effective treatments are limited<sup>3</sup>.

Recent advances in genomics and bioinformatics have provided unprecedented opportunities to elucidate the complex genetic architecture of vitiligo. Genome-wide association studies (GWAS) have identified numerous loci associated with vitiligo susceptibility, shedding light on the genetic underpinnings of the disease<sup>4,5</sup>. Other multiomics techniques, such as transcriptomics<sup>6</sup>, epigenomics<sup>7</sup>, and proteomics<sup>8</sup>, allow for a comprehensive investigation of vitiligo at multiple biological levels. Integrating these diverse data types can offer a holistic view of the molecular landscape of vitiligo, identifying key pathways and molecular targets involved in disease pathogenesis.

In this study, our aim was to utilize integrative multiomics data to uncover the genetic and molecular mechanisms underlying vitiligo and explore new therapeutic targets. We conducted a GWAS meta-analysis using the latest data from FinnGen, the GWAS Catalog, and the UK Biobank, revealing vitiligo-associated genetic loci through FUMA and MAGMA analyses<sup>9</sup>. Potential therapeutic targets were explored through a series of druggability analyses<sup>10</sup>. Additionally, we employed summary-based Mendelian randomization (SMR) methods<sup>11</sup>

<sup>1</sup>Department of Dermatology, Zhengzhou People's Hospital, Zhengzhou, Henan, China. <sup>2</sup>Department of Dermatology, First Affiliated Hospital of Zhengzhou University, No.1 Longhu Outer Ring Road, Jinshui District, Zhengzhou 450052, Henan, China. <sup>3</sup>Department of Dermatology, Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital, Nanjing, Jiangsu, China. <sup>4</sup>Department of Dermatology, The Sixth Medical Center of Chinese, PLA General Hospital, 6 Fucheng Road, Haidian District, Beijing 100048, China. <sup>5</sup>Zi-yue Dong, Ming-jie He and Yong-kai Yu have contributed equally to this work. ✉email: yrp7691@163.com; zja7691@163.com

combined with multiple eQTL and pQTL databases to identify pathogenic and protective genes for vitiligo, construct a transcription factor regulatory network, and validate findings via chromatin immunoprecipitation sequencing (ChIP-seq) data from the ENCODE database<sup>12</sup>.

To validate the genetically associated genes identified in the GWAS-meta analysis and the causal genes discovered through SMR analysis, we integrated microarray and single-cell RNA sequencing data<sup>13,14</sup> to investigate the differential expression of these genes in the blood and lesional skin of vitiligo patients. Our integrative approach highlights the role of the JAK/STAT pathway and its regulation of CTSS (Cathepsin S) in vitiligo, providing new insights into disease mechanisms and potential therapeutic targets.

## Materials and methods

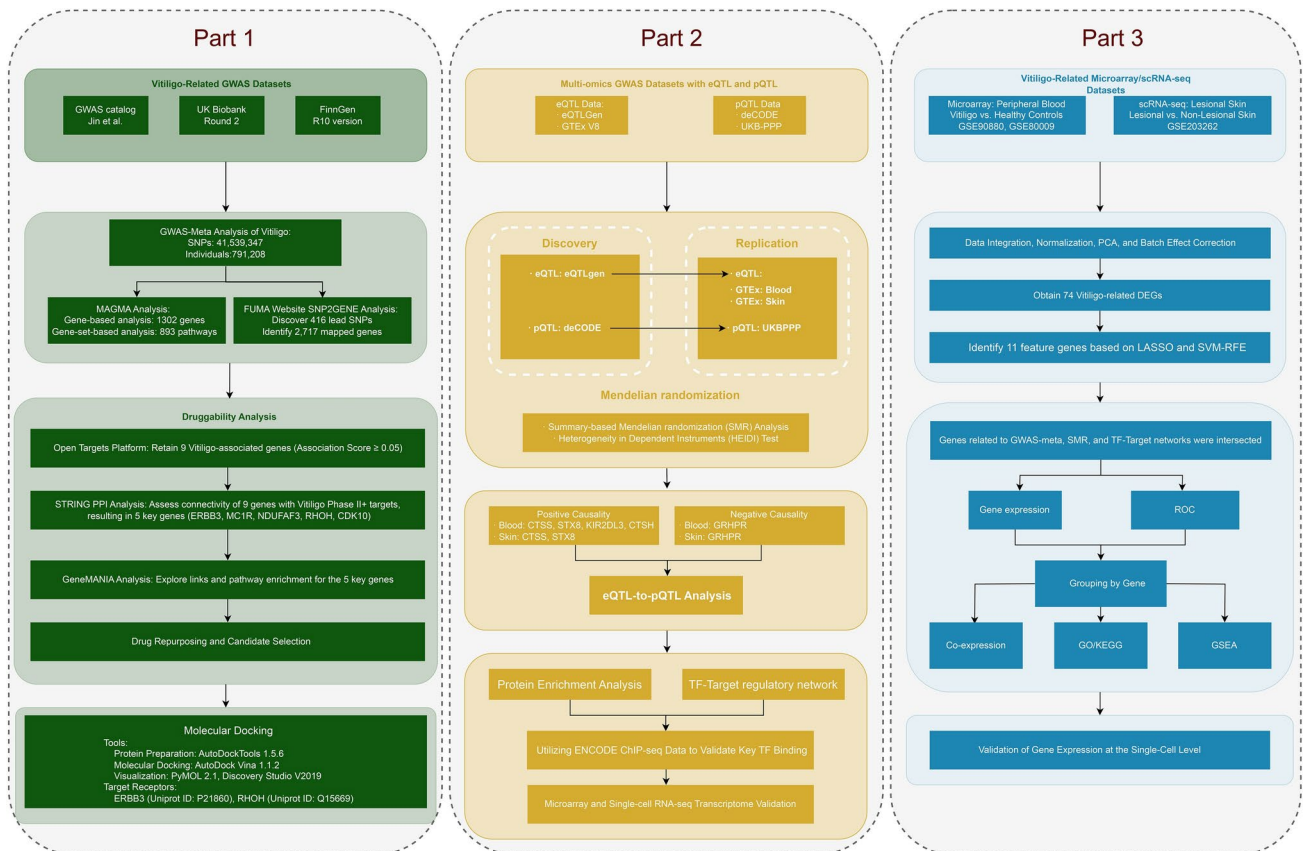
This study is divided into three interconnected parts, each focusing on different aspects of vitiligo research. The detailed design is shown in Fig. 1.

### Part 1: GWAS meta-analysis and druggability analysis

We integrated multiple vitiligo GWAS datasets, including those from the GWAS Catalog, UK Biobank, and FinnGen, comprising a total of 791,208 individuals and 41,539,347 SNPs. A GWAS meta-analysis was conducted to identify genetic loci associated with vitiligo. Subsequent annotation of SNPs to genes was performed via the FUMA website, followed by MAGMA analysis to identify genetically associated genes. These genes were then subjected to a series of druggability analyses. Part 1 included the following steps:

#### GWAS meta-analysis

We conducted a GWAS meta-analysis to identify genetic loci associated with vitiligo. METAL<sup>15</sup> ([https://genom.e.sph.umich.edu/wiki/METAL\\_Documentation](https://genom.e.sph.umich.edu/wiki/METAL_Documentation)) was used to perform the GWAS meta-analysis. A sample size weighted meta-analysis of Z scores (Stouffer's method) was employed. This method converts the observed effect direction and *P*-value in each study into a signed Z-score: a very negative Z-score indicates a small *P*-value and that the allele is associated with a lower disease risk or quantitative trait level. A large Z-score indicates a small *P*-value and that the allele is associated with a higher disease risk or quantitative trait level. The Z-scores for each allele were combined across samples via a weighted sum, where the weights are proportional to the square root of the sample size in each study<sup>16</sup>.



**Fig. 1.** Study design overview. Part 1: Genome-Wide Association Study (GWAS) Meta-Analysis and Druggability Analysis. Part 2: Causal Gene Identification and Regulatory Network Construction. Part 3: Transcriptomic Analysis and Mechanistic Exploration.

*FUMA analysis*

We used the Functional Mapping and Annotation (FUMA) platform (<https://fuma.ctglab.nl/snp2gene>) to annotate SNPs to genes, with all parameters set to the default thresholds provided by the FUMA website.

*MAGMA analysis*

We performed Multi-marker Analysis of Genomic Annotation (MAGMA) analysis to identify genetically associated genes. MAGMA (v1.10)<sup>17</sup> is designed for gene-based and gene-set-based association analyses. This approach enables the identification of functional genes or functional modules (such as regulatory pathways) associated with the trait of interest and facilitates the discovery of genes linked to multiple small-effect SNPs.

*Gene prioritization*

We used the Open Targets platform (<https://platform.opentargets.org/>) to obtain genes with a comprehensive score above 0.05. This platform integrates evidence from genetics, genomics, transcriptomics, drugs, animal models and scientific literature, facilitating the scoring and ranking of target–disease associations for the purpose of drug target discovery. These genes subsequently intersected with those identified through MAGMA analysis and FUMA annotation, leading to the identification of the most druggable target genes from the genetically associated genes.

*Protein-protein interaction (PPI) analysis*

We analyzed the interactions between the intersecting genes identified in "Gene prioritization" section and drug targets that have entered phase II or higher clinical trials for vitiligo via the STRING database (<https://string-db.org/>).

*Pathway enrichment analysis*

We considered genes that are linked to existing drug targets in the PPI network constructed via the STRING platform as key genes. These key genes may represent new potential drug targets. We used the GeneMANIA website (<http://genemania.org/>) to construct an enrichment analysis network.

*Drug target exploration*

For genes that are linked to existing drug targets but lack associated therapeutic agents, we explored relevant drugs via the Drug Signatures Database (DSigDB, <http://dsigdb.tanlab.org/DSigDBv1.0/>)<sup>18</sup> and the NetworkAnalyst platform<sup>19</sup>.

*Molecular docking*

We focused on the relevant drugs identified in "Drug target exploration" section, selecting those that have progressed to phase II clinical trials or higher as candidate therapeutic agents for vitiligo. Additionally, molecular docking studies included drugs widely used for vitiligo treatment, such as Betamethasone valerate, Clobetasol propionate, and Tacrolimus, to provide a basis for comparison. The protein structures were retrieved from the Protein Data Bank (PDB) (<https://www.rcsb.org/>), while the chemical structures of the drugs were obtained from PubChem<sup>20</sup> (<https://pubchem.ncbi.nlm.nih.gov/>). Proteins were pre-processed using AutoDockTools<sup>21</sup> 1.5.6 by removing crystallographic water molecules and adding polar hydrogen atoms to prepare them for docking. Docking simulations were performed with AutoDock Vina 1.1.2, using a grid box centered on the active domain of each protein. The grid box size was set to 30 Å × 30 Å × 30 Å with a grid point distance of 0.05 nm, allowing for sufficient molecular movement. Binding free energy (kcal/mol) was employed as the primary evaluation criterion, with lower binding energy indicating greater stability of the drug-receptor complex and a higher likelihood of interaction. Visualization of docking results was carried out using PyMOL 2.1 and Discovery Studio (DS) V2019 (Dassault Systèmes Biovia, San Diego, CA, USA). Detailed 2D interaction diagrams highlighted key interaction features, including hydrogen bonds and hydrophobic interactions. Comparisons of binding affinities between candidate drugs and widely used therapeutic agents were made to evaluate their relative efficacy and specificity.

**Part 2: causal gene identification and regulatory network construction***SMR analysis*

Using eQTL data from eQTLGen<sup>22</sup> and GTEx V8<sup>23</sup> and pQTL data from deCODE<sup>24</sup> and UKB-PPP<sup>25</sup>, we performed summary-data-based Mendelian randomization (SMR) analysis. The analysis was conducted via SMR 1.3.1 software, with SMR-format files utilized to examine the relationships between gene expression and protein expression quantitative trait loci (eQTLs and pQTLs) and vitiligo. We established discovery and validation sets to identify genes (proteins) causally related to vitiligo. SMR analysis helps to explore the mechanisms underlying vitiligo pathogenesis. The HEIDI test was primarily used to determine whether the phenotype mediated by gene SNPs was due to linkage disequilibrium effects<sup>26,27</sup>.

*eQTL-to-pQTL analysis*

We conducted eQTL-to-pQTL analysis<sup>28</sup> to investigate the relationship between gene expression and protein expression.

*Protein enrichment analysis and transcription factor prediction*

To further investigate the downstream mechanisms and upstream regulatory molecules of these causal genes, we conducted protein enrichment analysis and transcription factor prediction. Protein enrichment analysis was performed via the GeneMANIA website. Transcription factor prediction was conducted primarily through the TF-Target Finder website ([https://jingle.shinyapps.io/TF\\_Target\\_Finder/](https://jingle.shinyapps.io/TF_Target_Finder/)), developed by Wang Jin<sup>29</sup>, which

integrates multiple transcription factor databases to facilitate the prediction of transcription factors associated with target genes. For the causal genes identified through SMR analysis, their upstream transcription factors were required to be found in at least four databases. The transcription factor network was constructed via Cytoscape software (v3.10.1 <https://cytoscape.org/>).

#### *Validation of causal genes and transcription factor regulatory network*

To validate the causal genes and the transcription factor regulatory network, we performed transcriptomic analysis in Part 3 to verify gene expression levels. Additionally, we utilized transcription factor-related ChIP-seq data from the ENCODE database across multiple cell lines to validate the potential binding interactions between the predicted transcription factors and their target genes.

### **Part 3: transcriptomic analysis and mechanistic exploration**

#### *Dataset integration and preprocessing*

We integrated two peripheral blood microarray datasets from vitiligo patients obtained from the Gene Expression Omnibus (GEO) database (GSE90880, GSE80009). Standard preprocessing steps were applied, including normalization, batch effect correction, principal component analysis (PCA), and quality control, to ensure the reliability of the data.

#### *Differential expression and feature gene analysis*

We conducted differential gene expression analysis via the limma package to identify genes that were significantly upregulated or downregulated in vitiligo patients compared with healthy controls. Next, we applied least absolute shrinkage and selection operator (LASSO) and support vector machine-recursive feature elimination (SVM-RFE)<sup>30</sup> to identify feature genes specific to vitiligo patients by taking the intersection of the results from both methods.

#### *Validation and functional analysis of genes from part 1 and part 2*

We performed differential expression analysis on the genetically associated genes, causal genes, and transcription factors identified in Parts 1 and 2. The results were used to construct ROC curves. The final target genes were divided into high- and low-expression groups on the basis of their median expression levels, and subsequent coexpression analysis, GO/KEGG enrichment analysis, and GSEA were subsequently conducted.

#### *Single-cell level analysis*

For single-cell RNA sequencing analysis, we used data from vitiligo lesions in the GSE203262 dataset. The cells were filtered on the basis of the following criteria: the number of genes detected per cell was greater than 300 and less than 7000, the UMI count per cell was greater than 1000, and the largest 3% of the cells were excluded. Additionally, cells whose mitochondrial gene expression was greater than 10% and whose red blood cell gene expression was greater than 3% of the total gene expression were removed. Batch effects were corrected via the canonical correlation analysis (CCA) method, which primarily followed the Seurat V5 package in R<sup>31</sup>.

## **Results**

### **Results of part 1: GWAS meta-analysis and druggability analysis**

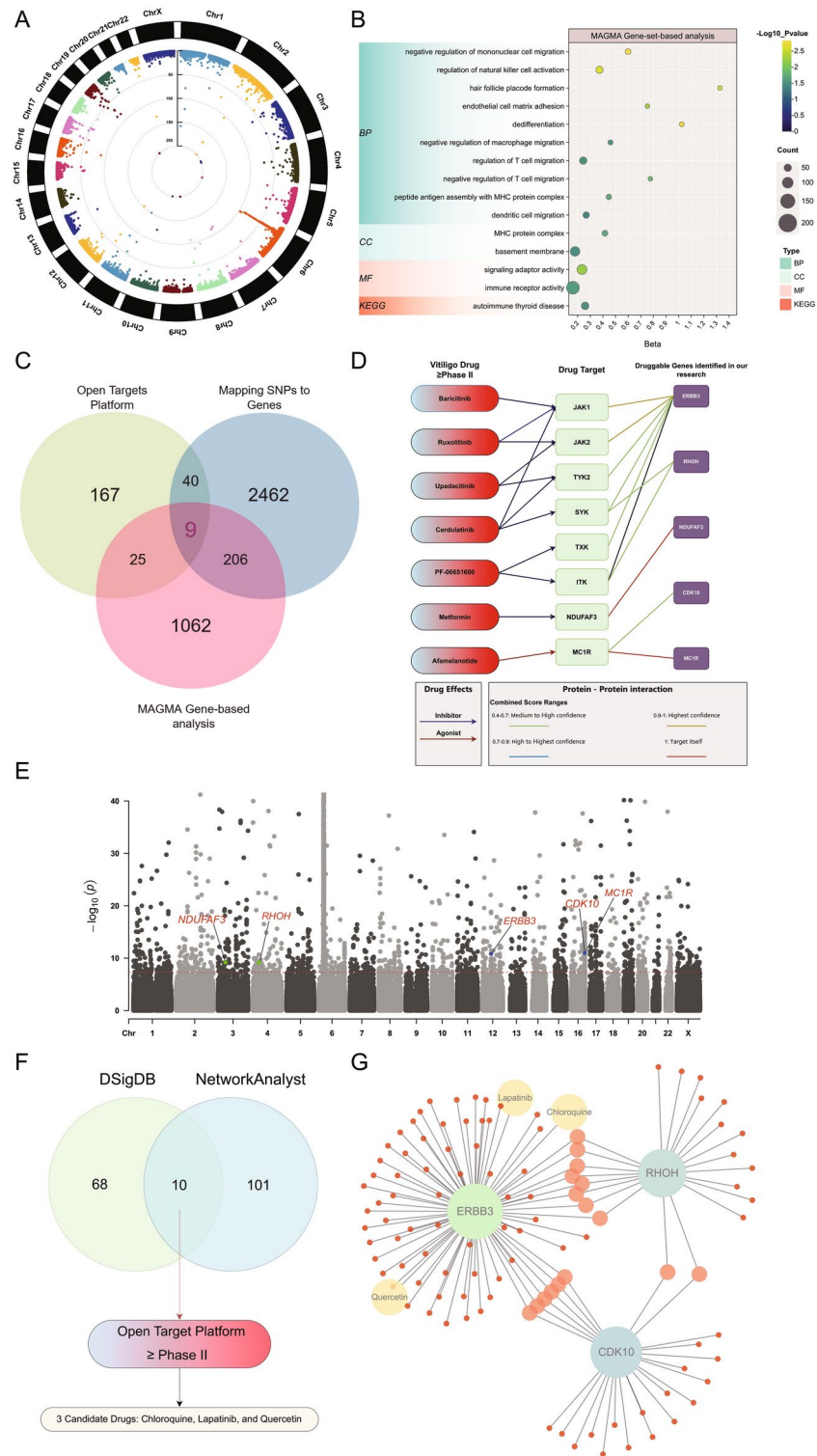
In Part 1 of our study, we conducted a comprehensive GWAS meta-analysis integrating datasets from the GWAS Catalog, UK Biobank, and FinnGen, encompassing a total of 791,208 individuals and 41,539,347 SNPs. The results of the meta-analysis revealed 9,558 SNPs strongly associated with vitiligo ( $P < 5e-08$ ), as shown in Fig. 2A. Using the FUMA platform with default parameters, we annotated these SNPs, resulting in 406 lead SNPs (Supplementary Fig. 1) and 2717 associated genes.

Next, we performed a genome-wide MAGMA analysis, identifying 1302 genetically associated genes and 893 pathways. We highlight 15 of these pathways in Fig. 2B. From the Open Targets Platform, we obtained 241 vitiligo-related targets with comprehensive scores above 0.05. By intersecting these targets with the genes identified through MAGMA and FUMA, we identified 9 genes with the highest druggability potential (Fig. 2C).

Further analysis of the protein-protein interactions (PPIs) of these 9 druggable genes with existing vitiligo treatment targets via the STRING database revealed 5 key genes that interact with current therapeutic targets, underscoring their potential as new drug targets (Supplementary Fig. 1). Figure 2D illustrates the relationships between these 5 key genes, existing vitiligo drugs, and their respective targets. Notably, the close association between ERBB3 and JAK1/2 suggests that any therapeutic benefit of tyrosine kinase inhibitors may potentially be derived from their additional impact on ERBB3. These genes are mapped onto the genome, and their locations are highlighted in the Manhattan plot Fig. 2E.

We then conducted pathway enrichment analysis via GeneMANIA, focusing on the 5 identified druggable genes. The resulting interaction network and enrichment analysis are depicted in Supplementary Fig. 3. This analysis highlighted potential pathways and processes in which these genes may be involved, further supporting their relevance in vitiligo pathology.

Notably, among the five druggable genes, the MC1R-targeting drug Afamelanotide appears to function more as an adjuvant therapy for vitiligo rather than as a standalone treatment option. Meanwhile, the NDUFAF3-targeting drug metformin is currently undergoing a phase 2 clinical trial, but no data have been published yet to support its efficacy in treating vitiligo. Therefore, our focus shifts to the candidate genes ERBB3 and RHOH, as these genes show the strongest associations with existing drug targets. To identify potential drugs targeting these genes, we used DSigDB and NetworkAnalyst to intersect and identify related drugs. We subsequently refined these findings via the Open Targets Platform and shortlisted three candidate drugs (Fig. 2F). Table 1 presents the relationships between the identified candidate drugs and their target genes based on DSigDB. One of the



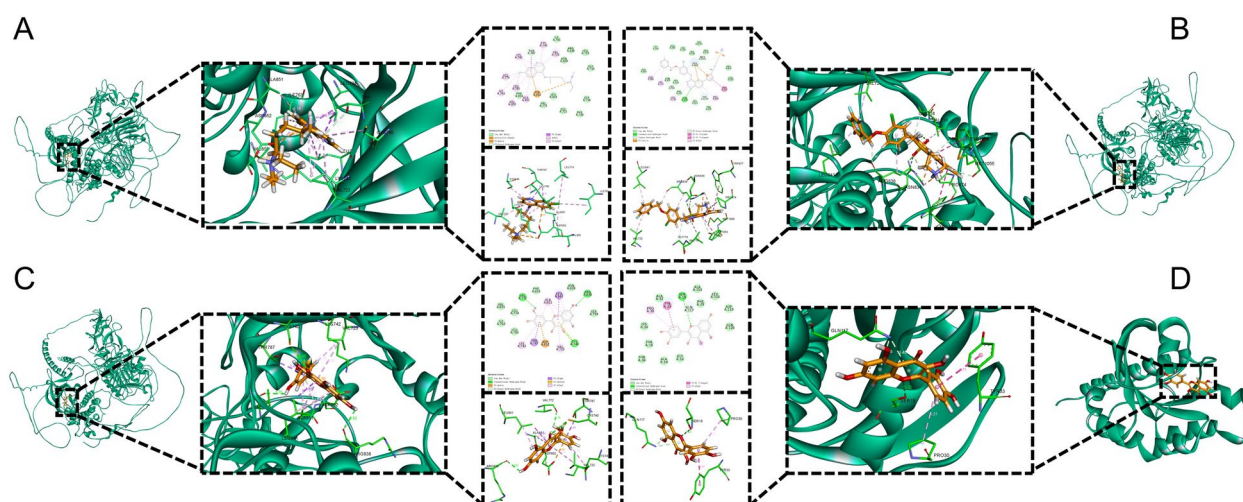
**Fig. 2.** Comprehensive Analysis of Genome-Wide Association Study (GWAS) Meta-Analysis and Druggability. (A) Genome-wide Manhattan plot illustrating the distribution of single nucleotide polymorphism (SNP) associations across all chromosomes. (B) Gene set enrichment analysis results from Multi-marker Analysis of Genomic Annotation (MAGMA), highlighting pathways associated with vitiligo. (C) Venn diagram showing the intersection of genes identified from MAGMA gene-based analysis, the Open Targets Platform, and Functional Mapping and Annotation (FUMA) Mapping SNPs to Genes. (D) Schematic representation of the relationship between existing vitiligo drugs, their targets, and the five newly identified druggable genes. (E) Locations of the five druggable genes on the genome-wide Manhattan plot. (F) Identification of related drugs for the five druggable genes using Drug Signatures Database (DSigDB) and NetworkAnalyst, followed by further screening through the Open Targets Platform to obtain three candidate drugs. (G) NetworkAnalyst visualization of the relationships between the three candidate drugs and the three candidate genes.

Drug names	P-value	Adjusted P-value	Genes
Chloroquine	0.005	0.029	ERBB3
Lapatinib	0.002	0.029	ERBB3
Quercetin	0.010	0.034	RHOH

**Table 1.** Candidate drugs predicted using DSigDB

Target	Uniprot ID	Drug	PubChem ID	Binding energy (kcal/mol)
ERBB3	P21860	Lapatinib	208908	-9.036
ERBB3	P21860	Quercetin	5280343	-8.232
ERBB3	P21860	Chloroquine	2719	-6.439
RHOH	Q15669	Quercetin	5280343	-6.425
ERBB3	P21860	Betamethasone Valerate	16533	-6.809
ERBB3	P21860	Clobetasol Propionate	32798	-7.380
ERBB3	P21860	Tacrolimus	445643	-16.867
RHOH	Q15669	Betamethasone Valerate	16533	-6.098
RHOH	Q15669	Clobetasol Propionate	32798	-5.717
RHOH	Q15669	Tacrolimus	445643	-11.908

**Table 2.** Docking results of available proteins with candidate small molecules and widely used vitiligo drugs.



**Fig. 3.** Docking results of the target drugs with their respective proteins. (A) Docking results of ERBB3 with Chloroquine. (B) Docking results of ERBB3 with Lapatinib. (C) Docking results of ERBB3 with Quercetin. (D) Docking results of RHOH with Quercetin. The figure includes three-dimensional (3D) structures of the target proteins and their interactions with the drugs, as well as 2D interaction diagrams highlighting key bonds and interaction sites. ERBB3: Receptor tyrosine-protein kinase erbB-3; RHOH: Rho-related GTP-binding protein RhoH.

candidate drugs, quercetin, was found to interact with the candidate target RHOH in the DSigDB database, whereas NetworkAnalyst indicated an interaction with the candidate target ERBB3 Fig. 2G.

Finally, to validate the effectiveness of these candidate drugs, we performed molecular docking studies, focusing on three candidate drugs: Lapatinib, Quercetin, and Chloroquine, which dock with ERBB3 and RHOH. The docking results are detailed in Table 2 and Fig. 3. The docking results demonstrated that Tacrolimus exhibited the strongest binding affinity to both ERBB3 and RHOH, with binding energies of  $-16.867$  kcal/mol and  $-11.908$  kcal/mol, respectively. These values were significantly lower than those of other tested compounds, highlighting Tacrolimus as a potential strong binder for these targets. Among the candidate drugs identified in this study, Lapatinib showed a notable binding affinity to ERBB3 ( $-9.036$  kcal/mol), outperforming Quercetin ( $-8.232$  kcal/mol) and Chloroquine ( $-6.439$  kcal/mol). Similarly, Quercetin exhibited the highest binding affinity to RHOH ( $-6.425$  kcal/mol) among the studied compounds, marginally surpassing Betamethasone Valerate ( $-6.098$  kcal/mol) and Clobetasol Propionate ( $-5.717$  kcal/mol). Overall, the results suggest that the candidate small molecules, particularly Lapatinib and Quercetin, exhibit promising binding capabilities when

compared to some conventional vitiligo treatments. Tacrolimus, however, stands out as the strongest binder across both targets, potentially due to its unique structural and pharmacological properties.

Detailed information on the genes identified through MAGMA analysis, genes annotated by FUMA, and vitiligo-related genes obtained from the Opentarget platform can be found in the Supplementary Tables S2–S4.

### Results of part 2: causal gene identification and regulatory network construction

In Part 2, we utilized SMR analysis to identify causal genes related to vitiligo. The SMR analysis, which used eQTL data from eQTLGen and GTEx V8, as well as pQTL data from deCODE and UKB-PPP, is summarized in Fig. 4. In the pQTL discovery set (deCODE database), we identified 55 causal proteins (Fig. 4A), of which 23 proteins remained causal in the pQTL validation set (UKB-PPP database) (Fig. 4B). Furthermore, the eQTL discovery set (eQTLGen) revealed that 9 out of the 23 pQTL causal proteins also had consistent causal relationships at the eQTL level (Fig. 4C).

Cross-tissue validation using the eQTL validation set (GTEx V8) in blood and skin confirmed the causal relationships of GRHPR, CTSH, CTSS, STX8, and KIR2DL3 (Fig. 4D,E). Notably, CTSS, STX8, and GRHPR exhibited causal relationships with vitiligo in both the blood and skin. Figure 4F–K display scatter plots and chromosome locus plots for CTSS, STX8, and GRHPR (using GTEx skin eQTL data as examples). The HEIDI test results for these analyses showed *P*-values greater than 0.01, supporting the robustness of our findings.

The eQTL-to-pQTL analysis results, detailed in Table 3, indicated a positive correlation between gene and protein expression for these five genes. We constructed a protein–protein interaction network for these four pathogenic genes (proteins) via GeneMANIA to explore shared pathogenic pathways (Fig. 4L).

Additionally, we predicted upstream regulatory transcription factors (TFs) via multiple databases and constructed a TF network, revealing that STAT1, STAT3, and CTCF are key TFs that regulate STX8, CTSH, and CTSS (Fig. 4M). Validation with ENCODE ChIP-seq data confirmed that STAT1 and STAT3 bind to the promoter region of CTSS and the enhancer region of CTSH, whereas CTCF binds to the promoter and enhancer regions of CTSH and STX8. Figure 4N shows peak heights in Integrative Genomics Viewer (IGV) software (v2.17.3 <https://software.broadinstitute.org/software/igv/>) with Autoscale, and the global view from IGV is provided in the Supplementary Fig. 4.

The results from the SMR analysis using data from deCODE, UKB-PPP, eQTLGen, and GTEx V8 are provided in the Supplementary Tables S5–S8.

### Results of part 3: transcriptomic analysis and mechanistic exploration

First, we combined and preprocessed the peripheral blood datasets GSE80009 and GSE90880 for vitiligo patients and healthy controls. The images before and after batch correction and PCA can be found in the Supplementary Figs. 5 and 6.

Figure 5A shows the heatmap of differentially expressed genes between vitiligo patients and healthy controls. We performed LASSO regression analysis to identify 13 feature genes (Fig. 5B,C) and SVM-RFE analysis to identify 24 feature genes (Fig. 5D,E). By intersecting these results via a Venn diagram, we identified 11 feature genes: LILRB4, PPP2R4, PI3, LCN2, MICB, RAPIGAP2, COL13A1, CTSS, STAT3, UTY, and F2R (Fig. 5F).

Among these genes, CTSS was identified as a causal gene in Part 2, which was validated through multilevel cross-tissue analysis, and STAT3 was identified as an upstream regulatory transcription factor. Additionally, these two genes are associated with vitiligo at the genome-wide level: CTSS was one of the 1302 genes with a MAGMA analysis *P*-value less than 0.05 ( $P=0.003$ ), and STAT3 was one of the 2717 genes annotated through FUMA analysis.

Figure 5G,H shows a violin plot illustrating the differential expression of CTSS and its corresponding transcription factor STAT3. The ROC curves for CTSS and STAT3 (Fig. 5I,J) had AUC values greater than 0.85, indicating good diagnostic value for vitiligo.

Figure 5K presents a heatmap of single-gene grouping for CTSS, which shows more significant differences than those in Fig. 5A. The coexpression heatmap of CTSS with other genes (Fig. 5L) suggested that CTSS and STAT3 were coexpressed in the peripheral blood.

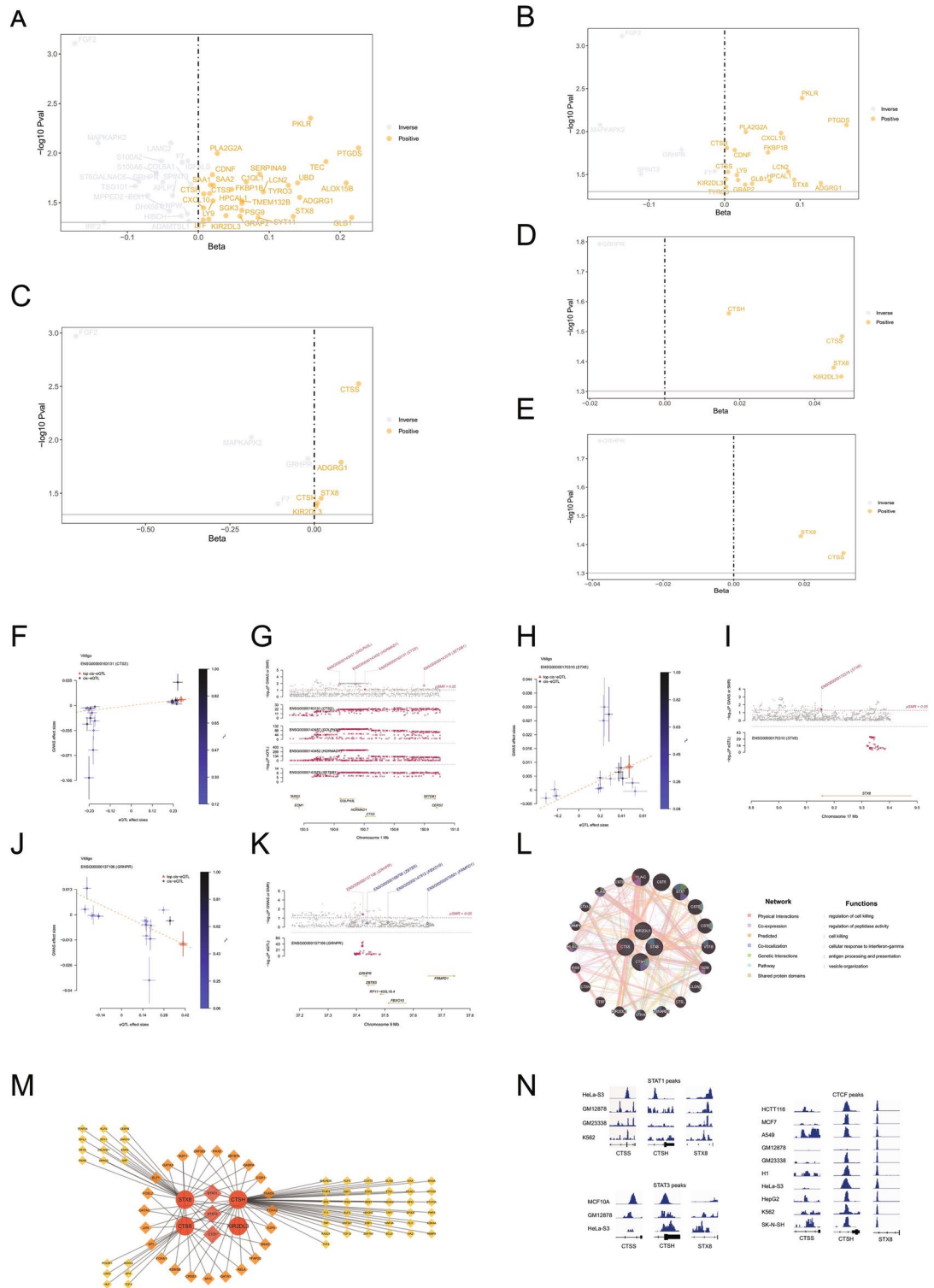
Further analysis included GO/KEGG enrichment analysis of the CTSS single-gene groups, which revealed 15 significant pathways (Fig. 5M). GSEA of the gene sets enriched in the high CTSS expression group (Fig. 5N) revealed that the most significant pathway was apoptosis, with the JAK/STAT signaling pathway among the top 5 pathways.

For single-cell RNA sequencing data analysis of vitiligo lesions (GSE203262), we focused on the differential expression of CTSS and its related transcription factors to further dissect potential mechanisms involved in vitiligo lesions. Figure 6A shows the UMAP plot of cell clustering, and Fig. 6C presents violin plots for STAT1, STAT2, STAT3, IFNGR1, IFNGR2, and CTSS, highlighting their differential expression across different cell types. We discovered that CTSS had the highest expression in dendritic cells. Compared with those in nonlesional areas, CTSS and STAT2 were highly expressed in melanocytes from lesional areas, whereas IFNGR1 and STAT3 were highly expressed in dendritic cells from lesional areas.

The standardized expression matrix after batch correction of vitiligo peripheral blood samples, the gene lists filtered by LASSO and SVM-RFE, and the ROC curves of the characteristic genes are available in the Supplementary Tables S9–S11. Additionally, detailed information on cell markers and differentially expressed genes across cell types can be found in the Supplementary Tables S12–S14.

## Discussion

This comprehensive multi-omics/multiomics bioinformatics study offers a novel and in-depth perspective on the potential mechanisms and therapeutic targets of vitiligo. Our research integrates several critical analytical approaches, including GWAS-meta analysis, MAGMA and FUMA analyses, druggability analysis, SMR analysis,



ChIP-seq analysis, microarray data analysis, and single-cell transcriptomic analysis. This integrative approach has provided significant insights into the genetic and molecular underpinnings of vitiligo, paving the way for the identification of new therapeutic targets and diagnostic markers.

In 2021, two-thirds of FDA-approved drugs had genetic evidence supporting their efficacy<sup>32</sup>, underscoring the increased likelihood of success for genetically supported drug targets in clinical trials<sup>33</sup>. In Part 1 of our study, a comprehensive genome-wide meta-analysis identified nine target genes with significant druggable potential and a strong genetic association with vitiligo. Notably, five of these genes have important links to existing vitiligo targets. For example, Melanocortin-1 receptor (MC1R) and NDUFAF3 are associated with targeted therapies. Afamelanotide, an MC1R agonist, was demonstrated in a multicenter randomized clinical trial<sup>34</sup> that the combination of afamelanotide implants with NB-UV-B phototherapy led to faster and more significant repigmentation than NB-UV-B monotherapy alone.



**Fig. 4.** (A) Protein quantitative trait loci (pQTL) discovery set from the deCODE database identified 55 causal proteins. (B) pQTL validation set from the UK Biobank Proteomics Project (UKB-PPP) database confirmed 23 of these proteins as causal. (C) Expression quantitative trait loci (eQTL) discovery set from eQTLGen found 9 of the 23 pQTL causal proteins with consistent causal relationships at the eQTL level. (D, E) Cross-tissue validation using Genotype-Tissue Expression (GTEx) V8 in blood (D) and skin (E) confirmed the causal relationships of glyoxylate reductase/hydroxypyruvate reductase (GRHPR), cathepsin H (CTSH), cathepsin S (CTSS), syntaxin 8 (STX8), and killer cell immunoglobulin-like receptor 2DL3 (KIR2DL3) with vitiligo. (F, H, J) Scatter plots for CTSS, STX8, and GRHPR respectively, using GTEx skin eQTL data as examples. Each point represents a SNP, with the x-axis showing its effect size in eQTLs and the y-axis showing its effect size in GWAS studies. The red triangle indicates the most significant cis-eQTL SNP, while other SNPs are color-coded by their linkage disequilibrium ( $R^2$ ) with the top SNP (color mapping shown in the legend). (G, I, K) Chromosome locus plots for CTSS, STX8, and GRHPR, respectively. The top panel shows GWAS SNPs as grey circles along the chromosome region, while the diamonds represent gene expression probes. Maroon diamonds indicate probes passing the Summary-based Mendelian Randomization (SMR) threshold, navy diamonds indicate probes not passing the SMR threshold, solid diamonds indicate probes passing the Heterogeneity in Dependent Instruments (HEIDI) threshold, and hollow diamonds indicate probes not passing the HEIDI threshold. (L) Protein-protein interaction network for the four pathogenic genes (proteins) constructed using Gene Multiple Association Network Integration Algorithm (GeneMANIA). (M) Transcription factor (TF) prediction network identifying signal transducer and activator of transcription 1 (STAT1), STAT3, and CCCTC-binding factor (CTCF) as key TFs regulating STX8, CTSH, and CTSS. (N) Validation of TF binding using Encyclopedia of DNA Elements (ENCODE) chromatin immunoprecipitation sequencing (ChIP-seq) data showing peak heights in Integrative Genomics Viewer (IGV) software with Autoscale; global IGV view provided in the Supplementary Fig. 4.

Gene	Beta	se	P value
CTSS	0.590179	0.028878	7.79E-93
GRHPR	0.30415	0.01029	5.11E-192
CTSH	1.60615	0.022812	3.58E-104
STX8	0.150016	0.021597	3.75E-12
KIR2DL3	1.47521	0.23654	4.47E-10

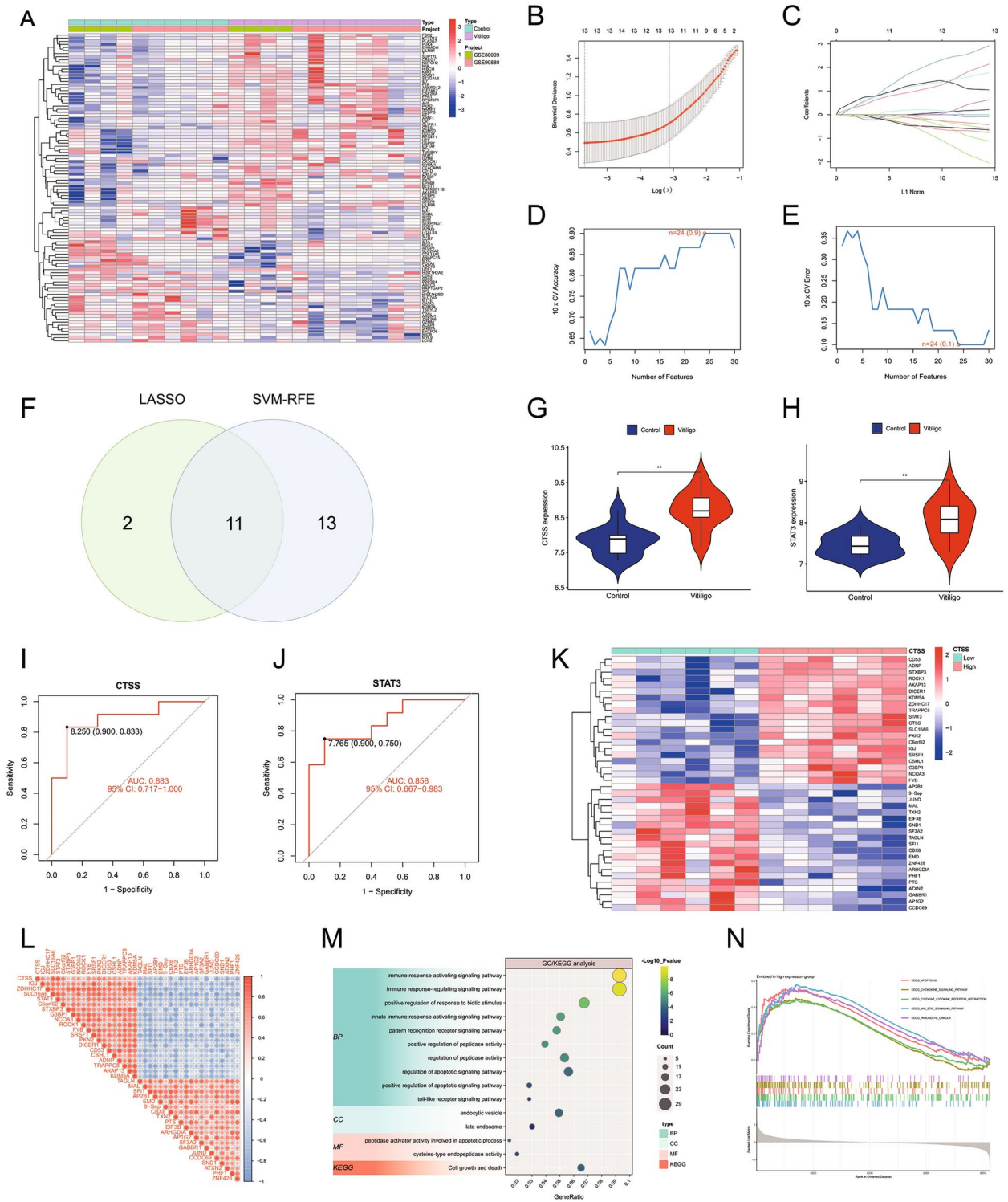
**Table 3.** Association between eQTL and pQTL of causally-related genes

In addition to MC1R, we identified other druggable targets, such as ERBB3 and RHOH, that currently lack targeted therapies. These targets are closely related to known vitiligo targets, such as JAK1, JAK2, TYK2, SYK, TXK, and ITK<sup>35</sup>. Through exploration of drug target databases and molecular docking studies, we identified lapatinib, quercetin, and chloroquine as the most promising therapeutic agents. Notably, quercetin targets both ERBB3 and RHOH. Previous studies have shown that quercetin inhibits  $H_2O_2$ -induced oxidative stress, thereby reducing melanocyte apoptosis<sup>36-38</sup>. In addition, the strong binding affinity of Tacrolimus to both ERBB3 and RHOH may explain its established clinical efficacy in vitiligo treatment. However, the candidate molecules, especially Lapatinib and Quercetin, demonstrated comparable or superior binding affinities to certain conventional drugs, indicating their potential as novel therapeutic agents for vitiligo. These results suggest that targeting ERBB3 and RHOH with these small molecules may offer a promising strategy for modulating the inflammation and immune dysregulation involved in vitiligo.

SMR is a powerful analytical method that leverages genetic variants as instrumental variables to infer causal relationships between gene expression and complex traits or diseases. This approach is particularly valuable in drug target discovery and the exploration of molecular mechanisms<sup>39,40</sup>. In our study, we identified five causal genes (proteins) through multilevel, cross-tissue SMR analysis. Subsequent chip analysis and single-cell transcriptomic validation revealed that only cathepsin S (CTSS) was differentially expressed and could effectively distinguish between vitiligo patients and healthy controls. These findings suggest that the CTSS plays a critical role in the pathogenesis of vitiligo.

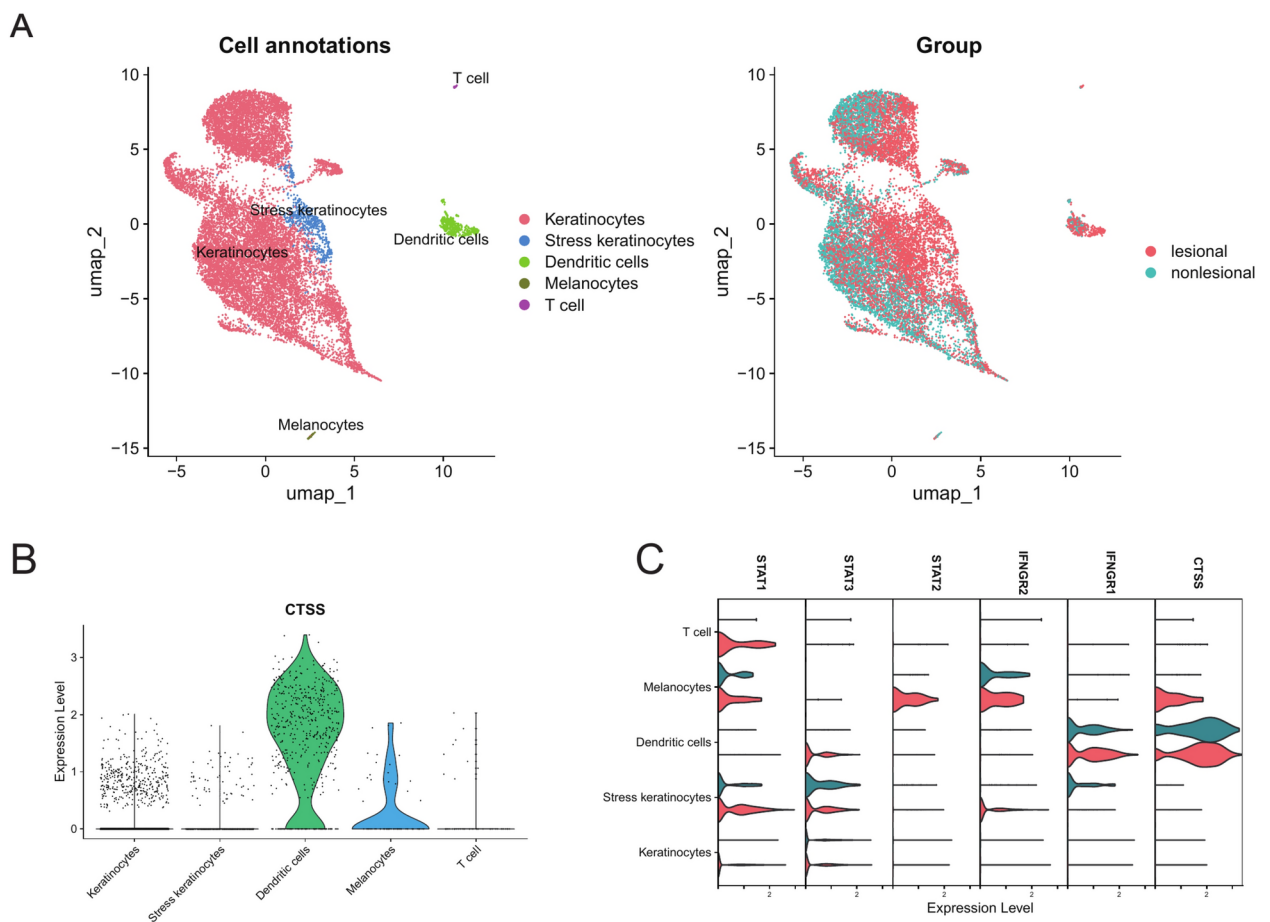
CTSS, a lysosomal cysteine protease belonging to the cathepsin family, is involved in various biological processes. Using the KEGG database<sup>41</sup> (<https://www.kegg.jp/kegg/>), we found that CTSS is part of the apoptosis pathway (hsa04210) and the antigen processing and presentation pathway (hsa04612), both of which are closely related to vitiligo development<sup>42,43</sup>. This finding aligns with our results, as shown in Fig. 2C's "peptide antigen assembly with MHC protein complex" pathway and Fig. 4L's "antigen processing and presentation" pathway. Notably, in Fig. 5N's GSEA analysis of the high CTSS expression group, "Apoptosis" emerged as the most significantly enriched pathway.

Anes et al.<sup>44</sup> systematically summarized the molecular functions of the cathepsin family and reported that CTSS is involved not only in the processing of MHC class II molecules but also in MHC class I-mediated antigen presentation. Exogenous antigens processed by CTSS in dendritic cells are ultimately presented to CD8+ T lymphocytes<sup>45</sup>. Our single-cell analysis of vitiligo lesions confirmed this finding, showing that CTSS is a marker for dendritic cells (Supplementary Table S14) and is differentially expressed in melanocytes, with nonlesional melanocytes not expressing CTSS Fig. 6.



In addition to exploring the downstream pathways mediated by CTSS, such as apoptosis and antigen presentation, we investigated the upstream mechanisms responsible for its differential expression. Using transcription factor prediction and ChIP-seq analysis, we identified STAT3 as the sole transcription factor that is differentially expressed both in peripheral blood and in lesional dendritic cells. STAT3 is also a characteristic gene of vitiligo, with AUC values for both STAT3 and CTSS molecules exceeding 0.85, indicating its potential diagnostic value. Additionally, we discovered that IFNGR1, an upstream molecule of STAT3<sup>46</sup>, is differentially expressed in lesional dendritic cells. While STAT1 and STAT2 were not found to be characteristic genes of vitiligo in peripheral blood, GSEA in the high CTSS expression group revealed enrichment of the JAK/STAT pathway. Furthermore, STAT1 and STAT2 were expressed at higher levels in lesional melanocytes than in nonlesional melanocytes. These results collectively suggest that the JAK/STAT pathway may serve as an upstream regulatory mechanism for CTSS, playing a crucial role in the pathogenesis of vitiligo.

**Fig. 5.** (A) Heatmap of differentially expressed genes (DEGs) between vitiligo patients and healthy controls. (B, C) Feature genes identified by Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis, showing 13 feature genes. (D, E) Feature genes identified by Support Vector Machine-Recursive Feature Elimination (SVM-RFE) analysis, showing 24 feature genes. (F) Venn diagram of intersecting feature genes from LASSO and SVM-RFE analyses, identifying 11 common feature genes: leukocyte immunoglobulin-like receptor subfamily B member 4 (LILRB4), protein phosphatase 2 regulatory subunit 4 (PPP2R4), peptidase inhibitor 3 (PI3), lipocalin 2 (LCN2), MHC class I polypeptide-related sequence B (MICB), RAP1 GTPase activating protein 2 (RAP1GAP2), collagen type XIII alpha 1 chain (COL13A1), CTSS, STAT3, ubiquitously transcribed tetratricopeptide repeat gene on the Y chromosome (UTY), and coagulation factor II receptor (F2R). (G, H) Violin plots showing the differential expression of CTSS and its transcription factor STAT3 between vitiligo patients and healthy controls. (I, J) Receiver operating characteristic (ROC) curves for CTSS and STAT3, indicating diagnostic value with area under the curve (AUC) values greater than 0.85. (K) Heatmap of single-gene grouping for CTSS, highlighting significant differences in gene expression. (L) Co-expression heatmap of CTSS with other genes, indicating co-expression with STAT3 in peripheral blood. (M) Gene Ontology (GO)/Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of CTSS single-gene grouping, showing the 15 significant pathways. (N) Gene Set Enrichment Analysis (GSEA) of gene sets enriched in the high CTSS expression group, highlighting apoptosis and Janus kinase/signal transducer and activator of transcription (JAK/STAT) signaling pathways among the top 5 pathways.



**Fig. 6.** (A) Uniform Manifold Approximation and Projection (UMAP) plot of cell clustering from single-cell RNA sequencing (scRNA-seq) data of vitiligo lesions (GSE203262), showing the distribution of different cell types. (B) CTSS exhibits the highest expression levels in dendritic cells, followed by melanocytes, making it a cellular marker for dendritic cells. (C) Violin plots for STAT1, STAT2, STAT3, interferon gamma receptor 1 (IFNGR1), IFNGR2, and CTSS, illustrating their differential expression across different cell types in vitiligo lesions. Detailed information on cell markers and differentially expressed genes across cell types can be found in the Supplementary Tables S12–S14.

While our comprehensive multiomics bioinformatics study provides significant insights into the potential mechanisms and therapeutic targets of vitiligo, several limitations must be acknowledged. First, the integration of various data types, including GWAS-meta analysis, MAGMA and FUMA analyses, druggability analysis, SMR analysis, ChIP-seq analysis, microarray data analysis, and single-cell transcriptomic analysis, introduces potential biases and technical variability. Each analytical method has inherent limitations, and the combination of these methods can sometimes amplify these issues. Second, our study relies heavily on publicly available databases and previously published data. The majority of the datasets used are derived from populations of European descent, potentially limiting the generalizability of our findings to other ethnic groups. Third, although we identified several promising therapeutic targets and pathways, our study is primarily computational and predictive. Experimental validation *in vitro* and *in vivo* is essential to confirm the biological relevance and therapeutic potential of these targets and pathways. Specifically, the therapeutic effects of quercetin on vitiligo, as well as the role of the JAK/STAT pathway in regulating CTSS involvement in antigen presentation and apoptosis pathways, require further experimental verification.

## Conclusion

In this comprehensive multiomics bioinformatics study, we have provided novel insights into the genetic and molecular mechanisms underlying vitiligo. Through the integration of GWAS-meta analysis, MAGMA and FUMA analyses, druggability analysis, SMR analysis, ChIP-seq analysis, microarray data analysis, and single-cell transcriptomic analysis, we identified several potential therapeutic targets and pathways, including CTSS and the JAK/STAT pathway. Our findings highlight the critical role of CTSS in the pathogenesis of vitiligo and suggest that the JAK/STAT pathway is a key regulatory mechanism. Additionally, we proposed quercetin as a potential therapeutic agent on the basis of molecular docking studies. While our study is primarily computational, it lays a solid foundation for future experimental research and clinical trials aimed at developing new treatments for vitiligo.

## Data availability

All original data sources are provided in Supplementary Table S1. An explanation of key bioinformatics terms used in this study is available in Supplementary Table S15. The protein and compound structures used for the molecular docking part of this study have been deposited in the Mendeley Data repository and can be accessed at <https://data.mendeley.com/datasets/pwfvn273j4/1>. All other data are included in this article and its supplementary information files. For further inquiries, please contact the corresponding author at [zja7691@163.com](mailto:zja7691@163.com).

Received: 15 September 2024; Accepted: 8 January 2025

Published online: 17 January 2025

## References

- Alikhan, A., Felsten, L. M., Daly, M. & Petronic-Rosic, V. Vitiligo: a comprehensive overview Part I. Introduction, epidemiology, quality of life, diagnosis, differential diagnosis, associations, histopathology, etiology, and work-up. *J. Am. Acad. Dermatol.* **65**(3), 473–491 (2011).
- Ezzedine, K., Eleftheriadou, V., Whitton, M. & van Geel, N. Vitiligo. *Lancet Lond Engl.* **386**(9988), 74–84 (2015).
- Rodrigues, M. et al. Current and emerging treatments for vitiligo. *J. Am. Acad. Dermatol.* **77**(1), 17–29 (2017).
- Jin, Y. et al. Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. *Nat. Genet.* **44**(6), 676–680 (2012).
- Jin, Y. et al. Genome-wide association studies of autoimmune vitiligo identify 23 new risk loci and highlight key pathways and regulatory variants. *Nat. Genet.* **48**(11), 1418–1424 (2016).
- Zhang, D. et al. Evidence of pyroptosis and ferroptosis extensively involved in autoimmune diseases at the single-cell transcriptome level. *J. Transl. Med.* **20**(1), 363 (2022).
- Pu, Y. et al. Transcriptome and differential methylation integration analysis identified important differential methylation annotation genes and functional epigenetic modules related to vitiligo. *Front. Immunol.* **12**, 587440 (2021).
- Berrevoet, D., Van Nieuwerburgh, F., Deforce, D. & Speeckaert, R. Proteomics data in vitiligo: A scoping review. *Front. Immunol.* **15**, 1387011 (2024).
- Gui, J. et al. A cross-tissue transcriptome-wide association study reveals novel susceptibility genes for migraine. *J. Headache Pain.* **25**(1), 94 (2024).
- Rasooly, D. et al. Genome-wide association analysis and Mendelian randomization proteomics identify drug targets for heart failure. *Nat. Commun.* **14**(1), 3826 (2023).
- Si, S., Liu, H., Xu, L. & Zhan, S. Identification of novel therapeutic targets for chronic kidney disease and kidney function by integrating multi-omics proteome with transcriptome. *Genome Med.* **16**(1), 84 (2024).
- ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**(7414), 57–74 (2012).
- Dey-Rao, R. & Sinha, A. A. Vitiligo blood transcriptomics provides new insights into disease mechanisms and identifies potential novel therapeutic targets. *BMC Genom.* **18**(1), 109 (2017).
- Shiu, J. et al. Multimodal analyses of vitiligo skin identify tissue characteristics of stable disease. *JCI Insight* **7**(13), e154585 (2022).
- Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinforma Oxf Engl.* **26**(17), 2190–2191 (2010).
- Kim, H. K. et al. Cross-disorder GWAS meta-analysis of endocannabinoid DNA variations in major depressive disorder, bipolar disorder, attention deficit hyperactivity disorder, autism spectrum disorder, and schizophrenia. *Psychiatry Res.* **330**, 115563 (2023).
- de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**(4), e1004219 (2015).
- Cao, Y., Yang, Y., Hu, Q. & Wei, G. Identification of potential drug targets for rheumatoid arthritis from genetic insights: A Mendelian randomization study. *J. Transl. Med.* **21**(1), 616 (2023).
- Xia, J., Gill, E. E. & Hancock, R. E. W. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat. Protoc.* **10**(6), 823–844 (2015).
- Wang, Y. et al. PubChem BioAssay: 2017 update. *Nucleic Acids Res.* **45**(D1), D955–D963 (2017).
- Morris, G. M., Huey, R. & Olson, A. J. Using AutoDock for ligand-receptor docking. *Curr. Protoc. Bioinforma* **24**(1), 8–14 (2008).

22. Vösa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**(9), 1300–1310 (2021).
23. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**(6509), 1318–1330 (2020).
24. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53**(12), 1712–1721 (2021).
25. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**(7982), 329–338 (2023).
26. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**(5), 481–487 (2016).
27. Wu, Y. et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* **9**(1), 918 (2018).
28. Chen, J. et al. Multi-omic insight into the molecular networks of mitochondrial dysfunction in the pathogenesis of inflammatory bowel disease. *EBioMedicine* **99**, 104934 (2024).
29. Wang, J. TTF: An R-based integrative tool for decoding human transcription factor–target interactions. *Biomolecules* **14**(7), 749. <https://doi.org/10.3390/biom14070749> (2024).
30. Zhou, B. et al. Identification and validation of CCR5 linking keloid with atopic dermatitis through comprehensive bioinformatics analysis and machine learning. *Front. Immunol.* **15**, 1309992 (2024).
31. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**(2), 293–304 (2024).
32. Ochoa, D. et al. Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. *Nat. Rev. Drug. Discov.* **21**(8), 551 (2022).
33. Trajanoska, K. et al. From target discovery to clinical drug development with human genetics. *Nature* **620**(7975), 737–745 (2023).
34. Lim, H. W. et al. Afamelanotide and narrowband UV-B phototherapy for the treatment of vitiligo: A randomized multicenter trial. *JAMA Dermatol.* **151**(1), 42–50 (2015).
35. Ma, J., Lyu, H., Huang, J. & Liu, B. Targeting of erbB3 receptor to overcome resistance in cancer treatment. *Mol. Cancer* **8**(13), 105 (2014).
36. Jeong, Y. M. et al. Cytoprotective effect of green tea extract and quercetin against hydrogen peroxide-induced oxidative stress. *Arch. Pharm. Res.* **28**(11), 1251–1256 (2005).
37. Guan, C. et al. Quercetin attenuates the effects of H<sub>2</sub>O<sub>2</sub> on endoplasmic reticulum morphology and tyrosinase export from the endoplasmic reticulum in melanocytes. *Mol. Med. Rep.* **11**(6), 4285–4290 (2015).
38. Li, X. et al. Combining network pharmacology, molecular docking, molecular dynamics simulation, and experimental verification to examine the efficacy and immunoregulation mechanism of FHB granules on vitiligo. *Front. Immunol.* **14**, 1194823 (2023).
39. Mao, R., Li, J. & Xiao, W. Identification of prospective aging drug targets via Mendelian randomization analysis. *Aging Cell* **4**, e14171 (2024).
40. Sun, J. et al. Identification of novel protein biomarkers and drug targets for colorectal cancer by integrating human plasma proteome with genome. *Genome Med.* **15**(1), 75 (2023).
41. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
42. Chen, J., Li, S. & Li, C. Mechanisms of melanocyte death in vitiligo. *Med. Res. Rev.* **41**(2), 1138–1166 (2021).
43. Xie, H. et al. Vitiligo: How do oxidative stress-induced autoantigens trigger autoimmunity?. *J. Dermatol. Sci.* **81**(1), 3–9 (2016).
44. Anes, E., Pires, D., Mandal, M. & Azevedo-Pereira, J. M. Spatial localization of cathepsins: Implications in immune activation and resolution during infections. *Front. Immunol.* **13**, 955407 (2022).
45. Shen, L., Sigal, L. J., Boes, M. & Rock, K. L. Important role of cathepsin S in generating peptides for TAP-independent MHC class I crosspresentation in vivo. *Immunity* **21**(2), 155–165 (2004).
46. Qing, Y. & Stark, G. R. Alternative activation of STAT1 and STAT3 in response to interferon-gamma. *J. Biol. Chem.* **279**(40), 41679–41685 (2004).

## Acknowledgments

We thank the researchers of the original genome-wide association studies (GWAS) for providing aggregated statistical data. We also express our gratitude to the Translational Medicine Center of the First Affiliated Hospital of Zhengzhou University for their support.

## Author contributions

Z-yD: Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing—original draft, Writing—review and editing. M-jH: Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing—original draft, Writing—review and editing. Y-kY: Methodology, Software, Supervision, Validation, Writing—original draft, Writing—review and editing. FW and P-yZ: Conceptualization, Formal Analysis, Investigation, Resources, Writing—review and editing. D-IR: Conceptualization, Formal Analysis, Resources, Software, Writing—review and editing. D-sF: Conceptualization, Methodology, Validation, Writing—review and editing. R-pY: Data curation, Investigation, Project administration, Writing—review and editing. QH: Conceptualization, Investigation, Software, Writing—review and editing. J-aZ: Writing—original draft, Writing—review and editing. All the authors read and approved the final manuscript.

## Funding

This research was supported by the Basic and Applied Research Program of the Zhengzhou Science and Technology Bureau (Grant No. 2024ZZJCYJ009).

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethics statement

The data used in this study were derived from previously published research, which obtained ethical approval from their respective committees. Therefore, no additional ethical permission was required for our study.

### Consent for publication

In the manuscript, consent was obtained from each individual for any form of personal data included.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-86134-4>.

**Correspondence** and requests for materials should be addressed to R.-p.Y. or J.-a.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025