Mathematical Biosciences and Engineering

MBE, 19(6): 5905–5924.
DOI: 10.3934/mbe.2022276
Received: 12 February 2022
Revised: 17 March 2022
Accepted: 18 March 2022
Published: 08 April 2022

http://www.aimspress.com/journal/MBE

Research article

Multiple datasets to explore the tumor microenvironment of cutaneous squamous cell carcinoma

Jiahua Xing[1,2], Muzi Chen[1,2] and Yan Han[1,*]

[1] Department of Plastic and Reconstructive Surgery, The First Medical Center, Chinese PLA General Hospital, Beijing 100853, China
[2] School of Medicine, Nankai University, Tianjin 300071, China

* **Correspondence:** Email: hanyanhyy10@163.com.

Abstract: *Background*: Cutaneous squamous cell carcinoma (cSCC) is one of the most frequent types of cutaneous cancer. The composition and heterogeneity of the tumor microenvironment significantly impact patient prognosis and the ability to practice precision therapy. However, no research has been conducted to examine the design of the tumor microenvironment and its interactions with cSCC. *Material and Methods*: We retrieved the datasets GSE42677 and GSE45164 from the GEO public database, integrated them, and analyzed them using the SVA method. We then screened the core genes using the WGCNA network and LASSO regression and checked the model's stability using the ROC curve. Finally, we performed enrichment and correlation analyses on the core genes. *Results*: We identified four genes as core cSCC genes: DTYMK, CDCA8, PTTG1 and MAD2L1, and discovered that RORA, RORB and RORC were the primary regulators in the gene set. The GO semantic similarity analysis results indicated that CDCA8 and PTTG1 were the two most essential genes among the four core genes. The results of correlation analysis demonstrated that PTTG1 and HLA-DMA, CDCA8 and HLA-DQB2 were significantly correlated. *Conclusions*: Examining the expression levels of four primary genes in cSCC aids in our understanding of the disease's pathophysiology. Additionally, the core genes were found to be highly related with immune regulatory genes, suggesting novel avenues for cSCC prevention and treatment.

Keywords: cutaneous squamous carcinoma; tumor microenvironment; tumor infiltrating immunocyte; bioinformatics; immune infiltration analysis

1. Introduction

As a non-melanoma skin cancer, cutaneous squamous cell carcinoma(cSCC) has a considerable influence on worldwide health systems. According to linked studies, non-melanoma skin cancer was the most frequent malignancy in men and women in 2017, with up to 7.7 million instances, of which 5.9 million cases were basal cell carcinoma and 1.8 million cases were cSCC [1]. In 2020, 1.2 million people develop cSCC, according to a recent study [2]. Even though cSCC has a low risk of lymphatic metastasis and a favorable prognosis, 2.1 to 5% of skin cancers advance to a metastatic state [3,4]. As the most frequent type of metastatic skin cancer, cSCC has a terrible prognosis at the late stage [5], accounting for 20% of cutaneous carcinoma-related deaths [4]. CSCC is a type of keratin-forming cell tumor that accounts for around 20% of all keratin-forming cell malignancies [6]. Long-term exposure to ultraviolet radiation, irregular circadian rhythms, smoking and alcohol misuse, and being overweight are frequent risk factors for skin cancer [7,8]. Additionally, it may develop due to various skin conditions such as scars, burn wounds, discoid erythema, or hypertrophic lichen planus [9–12]. Excision surgery is an excellent therapeutic option for extensive squamous skin cancers [13]. Due to its propensity for metastasis and poor prognosis in its advanced stages, we must continue investigating its pathophysiology and developing more effective treatments.

Further research revealed that squamous cell carcinomas, such as those found in the esophagus, the lung, the head and neck, and other locations, are complex tissue composed of various distinct cells [14–16]. We believe that cSCC is similar in tissue complexity to other types of squamous cell carcinoma. That is, squamous tumor tissue contains cancer cells and a diverse array of other cellular components involved in inflammation, angiogenesis, and the immune response, forming a complex network of cellular regulation [17].

The tumor microenvironment is constituted of tumor cells, resident and recruited host cells (cancer-associated stromal cells and immune cells), as well as the secretory products of the corresponding cells (e.g., cytokines and chemokines) and extracellular matrix components (ECM). The tumor microenvironment serves as the "soil" for tumor survival, and tumor growth and metastasis are highly dependent on the surrounding environment, analogous to the "seed-soil" relationship [18]. Angiogenesis is required for cancer cells to grow and metastasize in the tumor microenvironment [19,20]. The tumor microenvironment contains a diverse array of immune cells, which play a critical role in anticancer immunotherapy. In-depth examination of tumor-infiltrating immune cells can generate novel therapeutic avenues and concepts and produce particular anti-tumor immune responses in specific patient populations [21,22]. Thus, tumor cells and the tumor microenvironment complement one another, and the heterogeneity of the tumor microenvironment and the variable response rates of tumor patients to immunotherapy are inextricably linked. We can improve clinical treatment and achieve precision medicine by typing the tumor immune microenvironment.

We now have a better grasp of the molecular mechanisms and expression of oncogenes in the tumor microenvironment because of advances in research methodologies in recent years. Yan's study, for example, used single-cell RNA-sequencing technology(scRNA-seq) to investigate cSCC at the transcriptome level, collecting 350 single-cell transcriptome profiles from skin squamous cancer cells and normal tissue cells to investigate the fundamental biological pathways at play. Members of the S100 gene family, SPRR gene family, and FABP5 gene family are upregulated in cSCC cells, suggesting that S100A9 and FABP5 could be potential therapeutic targets for cSCC [23]. Andrew's

research is the first to combine geographic mapping of spatial transcriptomics (ST) and multiplexed ion beam imaging (MIBI) with scRNA-seq to create the most comprehensive single-cell transcriptional atlas of cell subpopulations inside cSCC to date, and to analyze the ecological niche in which these cells live. Patients with high expression of tumor-specific keratinocyte (TSK) markers ITGB1 and PLAU had significantly poorer progression-free survival after PD-1 inhibitor treatment, according to this study [24].

Given the importance of the immune microenvironment in carcinogenesis and the lack of information about cSCC-associated immune microenvironment, we investigated the interactions between the two and the mechanisms involved. We first investigated the molecular network mechanisms underlying the tumor microenvironment of cSCC using multiple datasets, performing immune cell infiltration analysis, WGCNA analysis, GO and KEGG functional analysis, and developing predictive models for the tumor microenvironment using lasso regression. We used GSEA analysis to compare the route differences between the high and low expression groups. Then we used enrichment analysis and Motif-TF annotation to identify transcription factors that can alter core gene regulation. Additionally, we screened essential genes based on their GO semantic similarity, did a differential analysis on immune-regulated genes, and investigated the relationship between core genes and immune regulation.

2. Material and methods

2.1. Data download

We retrieved the GSE42677 Series Matrix File data file from the NCBI GEO public database, along with the annotated file GPL571, and included the expression profile data for 20 patient groups, including the regular group (n = 10) and the tumor group (n = 10). Additionally, we obtained the GSE45164 Series Matrix File data file, annotated with GPL571, which contained the expression profiles of 13 patient groups, comprising normal (n = 3) and tumor (n = 10) groups. To adjust the microarray, we employed the SVA method. Then, we obtained the GSE53462 dataset, which was annotated using the platform GPL10558 and had ten distinct groups of transcriptome data, comprising a regular group (n = 5) and a tumor group (n = 5). We retrieved the GSE45216 dataset for subsequent validation, annotated on GPL570, containing 40 sets of transcriptome data, including AK group (n = 10) and SCC group (n = 30).

2.2. Immune cell infiltration analysis

The CIBERSORT method is a well-established technique for determining the immune cell types present in the microenvironment. The method is based on support vector regression, and deconvolution is used to assess the expression matrix of immune cell subtypes. It contains 547 biomarkers that distinguish 22 distinct subpopulations of human immune cells, including T-cell, B-cell, plasma cell, and myeloid cell. The CIBERSORT method was used to infer the 22 immune infiltrating cells' relative proportions and perform spearman correlation analysis on gene expression and resistant cell content in this study [25].

2.3. WGCNA analysis

By developing weighted gene co-expression networks, we identified co-expressed gene modules. We investigated the relationship between gene networks and phenotypes, as well as the core genes in the networks. We constructed a co-expression network for all genes in the GSE131761 dataset using the WGCNA-R package and selected genes with the top 5000 variances for further analysis using this technique, with the soft threshold set to 11. To evaluate network connectedness, we turned the weighted adjacency matrix into a topological overlap matrix (TOM) and used hierarchical clustering to generate the TOM matrix of the clustering tree structure. The different branches of the clustering tree correspond to distinct gene modules, and the various colors correspond to specific modules. We identify genes based on their weighted correlation coefficients and put genes with similar expression patterns into a module, splitting tens of thousands of genes into many modules based on their gene expression patterns.

2.4. GO and KEGG analysis

We used the Metascape database (www.metascape.org) for annotation and visualization, Gene Ontology (GO) analysis, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for specific genes to determine the biological activities and signaling pathways involved in illness development. We defined statistical significance as Min overlap $\geqslant$ 3.00 or $P \leqslant$ 0.01. Both GO/KEGG and Lasso prediction models used genes from the salmon module for 33 samples.

2.5. Predictive model construction

We picked the WGCNA modular genes most strongly associated with the trait and further developed prediction models using lasso regression. After combining the expression data for each gene, a scoring system was developed for each patient and weighted using the predicted regression coefficients in lasso regression analysis. We utilized the median score value as the cut-off point for the risk score formula and ROC curves to determine the model's predictive accuracy. We created a WGCNA network based on the patients' T-cell characteristics to investigate the critical regulatory genes in SCC. The function "sft$powerEstimate" was used to determine the soft threshold, which was set to 11.

2.6. GSEA analysis

GSEA analysis ranks genes based on their differential expression between two types of samples. It then determines whether the predetermined genes are enriched at the top or bottom of this ranking table. In this study, we used GSEA to compare the differential KEGG signaling pathway between the high and low expression groups to better understand the molecular mechanisms underlying the core genes in the two groups, with the number of substitutions set to 1000 and the type of substitution set to phenotype. The software used for GSEA analysis is GSEA4.0, and the background gene set is KEGG v7.5.symbol.gmt. GSEA analysis used all genes from the expression profile.

2.7. Regulatory network analysis of essential genes

In eukaryotes, transcription initiation is a highly complex process that frequently involves the cooperation of numerous protein components. Together, transcription factors and RNA polymerase II create a transcription initiation complex involved in the transcription initiation process. Transcription factors are classified into two types based on their mode of action. The first category includes universal transcription factors, which form a transcription initiation complex with RNA polymerase II before transcription can begin correctly. Cis-acting elements are sequences found in the flanking regions of genes that can influence gene expression. Cis-acting elements, which regulate gene expression, include promoters, enhancers, regulatory sequences, and inducible regions. The cis-acting element does not encode any protein. It merely supplies an action site for the trans-acting factor to bind with. The current study used the R package cisTarget, with the following parameters, rcistarget.hg19. motifdb. cisbpont.500bp, for the Gene-motif ranking database.

2.8. GO semantic similarity

We classified proteins by their functional similarity to their interacting proteins based on the similarity of GO semantics utilized for gene annotation. Correlation with gene expression profiles validated GO semantic similarity [26], establishing a basis for functional comparison of gene products. As a result, GO semantic similarity is widely used in bioinformatics applications such as protein-protein interaction analysis [27], pathway analysis [28], and gene function prediction [29]. In this section, we quantify the functional similarity of proteins. The term "functional similarity" refers to the geometric mean of the semantic similarity of GO terms in terms of cellular components (CC) and biological pathways (BP). It quantifies the strength of the relationship between each protein and its interacting proteins by considering both function and path. The GOSemSim software was used to determine the semantic similarity of interacting histones in CC and BP [30], which assumes the GO topology to perform more accurately [31]. The geometric mean of semantic similarity in CC and BP was used to evaluate functional similarity further.

2.9. Immunohistochemistry of key genes

To further validate our findings, we obtained immunohistochemical staining results of four key genes DTYMK, CDCA8, PTTG1 and MAD2L1 in normal skin tissue and tumor tissue on the Human Protein Atlas (https://www.proteinatlas.org/) website, enabling our core genes to be validated again after model validation.

2.10. Statistical analysis

The R programming language was used to conduct all statistical analyses in this study (version 4.0). All statistical tests were two-sided, and a significance level of $P < 0.05$ was considered statistically significant. The present analysis specifically involved two groups of patients (control and tumor groups) and did not perform multiple comparisons, therefore no corrected P values were required.

3. Results

We retrieved the GEO database's SCC-related datasets GSE42677 and GSE45164 and included 33 patients, 13 in the regular group and 20 in the malignant group. We use the SVA algorithm to fix the chips. The results indicated that the batch effect between chips was abolished when the SVA algorithm corrected for it (Figure 1A,B). The microenvironment mainly consists of fibroblasts, immune cells, extracellular matrix, numerous growth factors, inflammatory factors, and unique physicochemical properties, among others. The microenvironment significantly impacts the disease's diagnosis, survival outcome, and clinical treatment sensitivity. The association between the SCC dataset and immune infiltration was studied to understand better the underlying mechanisms affecting tumor progression. We depict a bar graph of the various types of immune cells found in each sample (Figure 2A). We created a heat map of immune cell correlations, highlighting positive correlations with red and negative correlations with blue, with darker colors suggesting stronger correlations (Figure 2B). Additionally, we discovered that when we compared the significance of microenvironment scores between tumor and regular groups, we found that B cells memory (P = 0.016), T cells regulatory (Tregs) (P = 0.001), Mast cells resting (P = 0.002), Mast cells activated (P = 0.02), and several other immune microenvironment factors were significantly different between groups (Figure 2C).

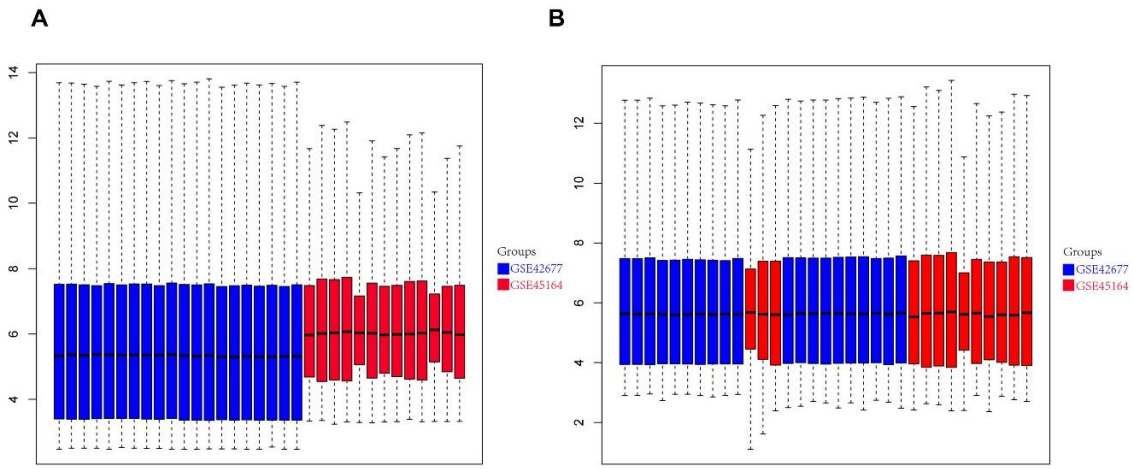


Figure 1. Removing batch effects between chips. (A) pre-corrected box line plots of GSE42677 and GSE45164; (B) corrected box line plots of GSE42677 and GSE45164. Blue group representatives GSE42677, red group representatives GSE45164.

We next created a WGCNA network and identified gene modules using the tom matrix, resulting in the identification of nine gene modules, including black (253), blue (482), greenyellow (865), grey (304), magenta (208) and purple (865). Modules are available in greenyellow (865), grey (304), magenta (208), purple (166), red (2502), salmon (102) and tan (118). We discovered the strongest association between the salmon module and SCC phenotypes by module-trait analysis (cor = 0.54, P = 0.001) (Figure 3). We performed enrichment analysis using the metascape database. We discovered that genes mainly involved in the following pathways, including mitotic cell cycle

process、microtubule cytoskeleton organization involved in mitosis、centrosome、Cell cycle、positive regulation of cell cycle (the top half of Figure 4A). The bottom half of Figure 4A is a cluster network of enriched pathways, where nodes sharing the same cluster are usually close to each other. The interplay of the salmon module's genes is seen in Figure 4B. The interaction network in Figure 4B was obtained from the string database to obtain the interactions between proteins and then visualized by cytoscape.

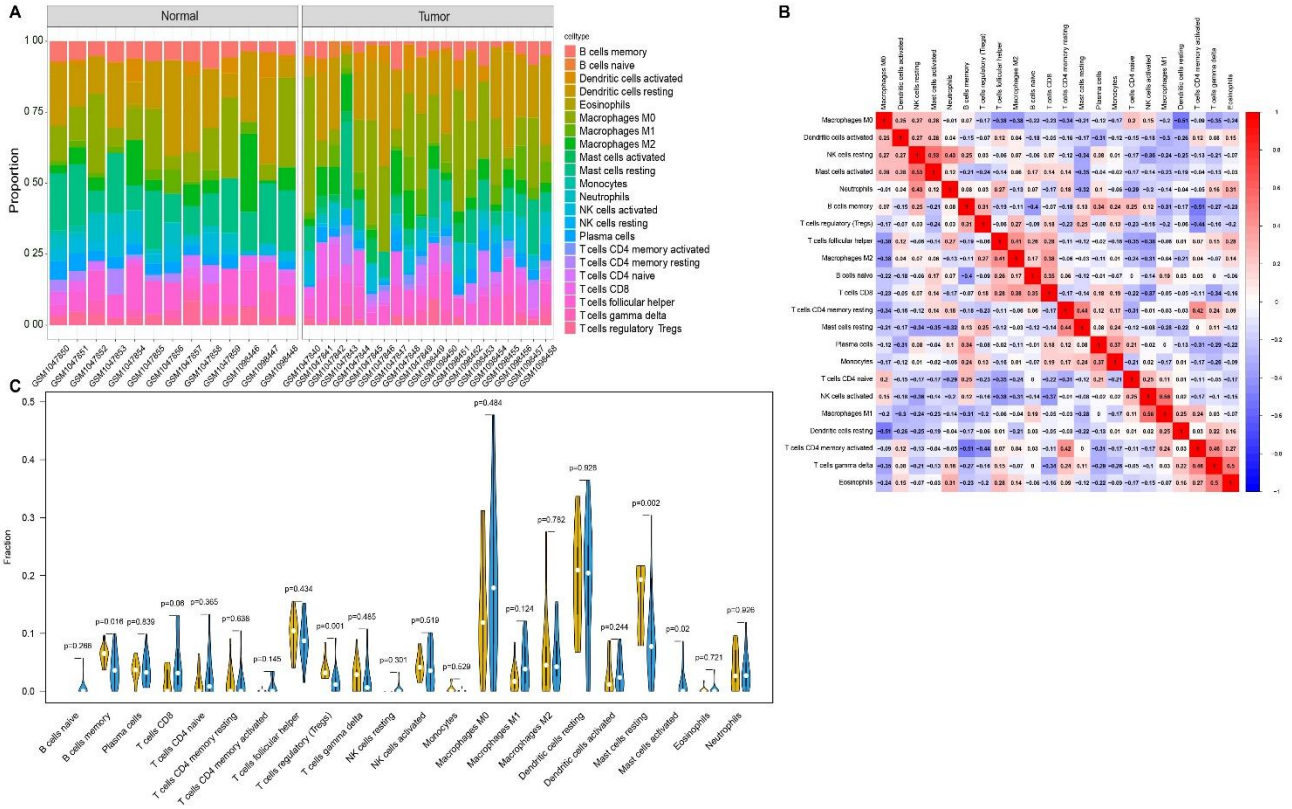


Figure 2. Microenvironmental immune infiltration analysis. (A) bar graph of different types of immune cells in each sample; (B) heat map of immune cell correlation; (C) differential analysis of infiltration of different kinds of immune cells in the tumor and regular groups.

To further characterize the salmon module gene set's essential genes, we employed the datasets GSE42677 and GSE45164 as training sets and GSE53462 and GSE45216 as validation sets. We used lasso regression to do feature screening. The results indicated that lasso regression revealed four genes as hallmark genes of SCC for further study. These four genes were DTYMK, CDCA8, PTTG1, and MAD2L1 (Figure 5A–C). Our study used the lasso algorithm to construct the prediction model. The results indicated that the model built using the four genes had a high diagnostic efficacy, with an area under the AUC curve of 0.9846. We subsequently tested the diagnostic model using the validation set, and the findings indicated that the model was relatively stable, with an area under the AUC curve of 0.8567 and 1, respectively (Figure 6A–C). Since the area under AUC curve greater than 0.8 can also be considered to have a relatively strong predictive power, our results suggest that the prediction model can predict not only the difference between normal vs SCC, but also AK vs SCC.

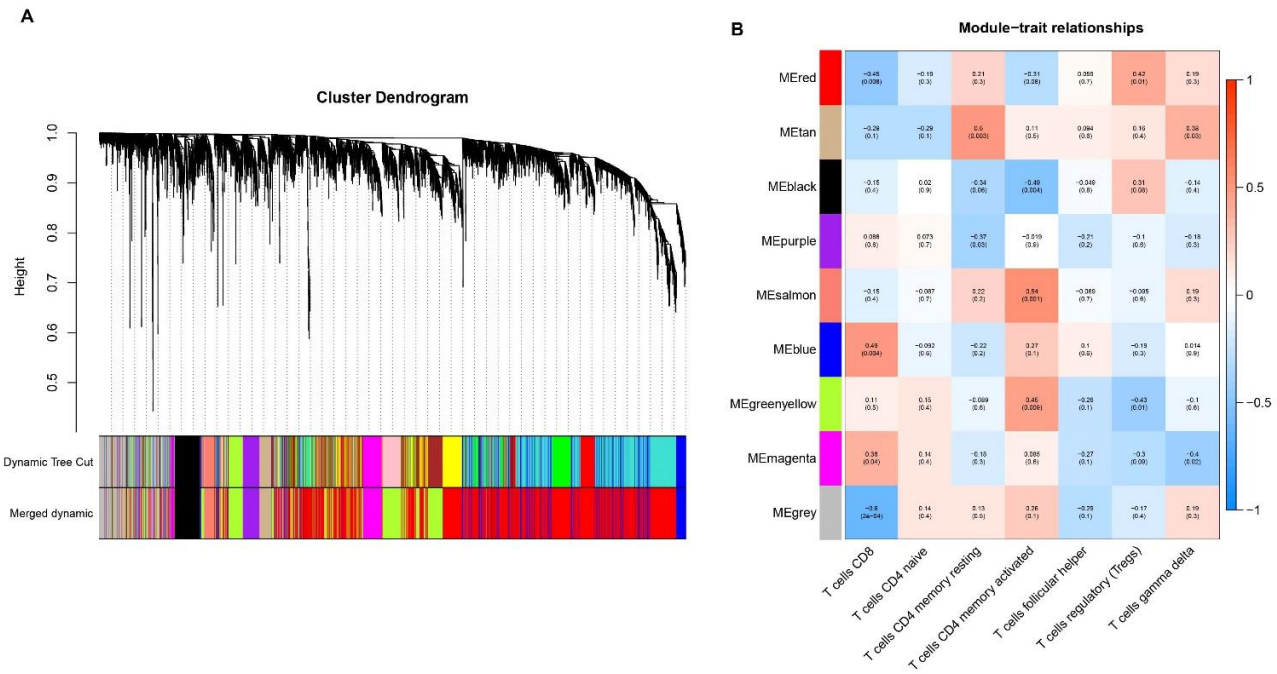


Figure 3. (A) A co-expression matrix was constructed to group genes into different modules based on their adjacency and similarity, and this was used to obtain a systematic clustering tree among genes; (B) Based on the tom matrix to detect gene modules, a total of nine gene modules were detected, namely black, blue, greenyellow, grey, magenta, purple, red, salmon, and tan. further exploring the correlation analysis between modules and traits, it was found that the salmon module had the highest correlation with the SCC phenotype (cor = 0.54, P = 0.001).

We then examined the individual signaling pathways enriched by the four core genes to determine the molecular mechanisms affecting SCC progression. GSEA analysis revealed considerable enrichment in some relevant pathways. The pathways with the highest CDCA8 expression levels were FRUCTOSE AND MANNOSE METABOLISM and CITRATE CYCLE TCA CYCLE. The paths with the highest levels of DTYMK expression were HUNTINGTONS DISEASE and KEGG PYRIMIDINE METABOLISM. MAD2L1 enrichment pathways include RNA DEGRADATION and CELL CYCLE. PTTG1 enrichment pathways are highly expressed in PARKINSONS DISEAS and PATHOGENIC ESCHERICHIA COLI INFECTION (Figure 7).

We analyzed four essential genes and discovered that they are regulated by various transcription factors and other standard processes. As a result, we used cumulative recovery curves, Motif-TF annotation, and a selection of essential genes to do enrichment analysis on these transcription factors. The research revealed that RORA, RORB, and RORC were the primary regulators of the gene set, with MOTIF being annotated as cisbp M5032. We identified two model genes with a normalized enrichment score (NES) of 7.9 enriched in this motif. We display all of the motifs and transcription factors associated with the model genes (Figure 8). Additionally, we screened essential genes based on their GO semantic similarity, and the results indicated that CDCA8 and PTTG1 were more critical across the network (Figure 9).

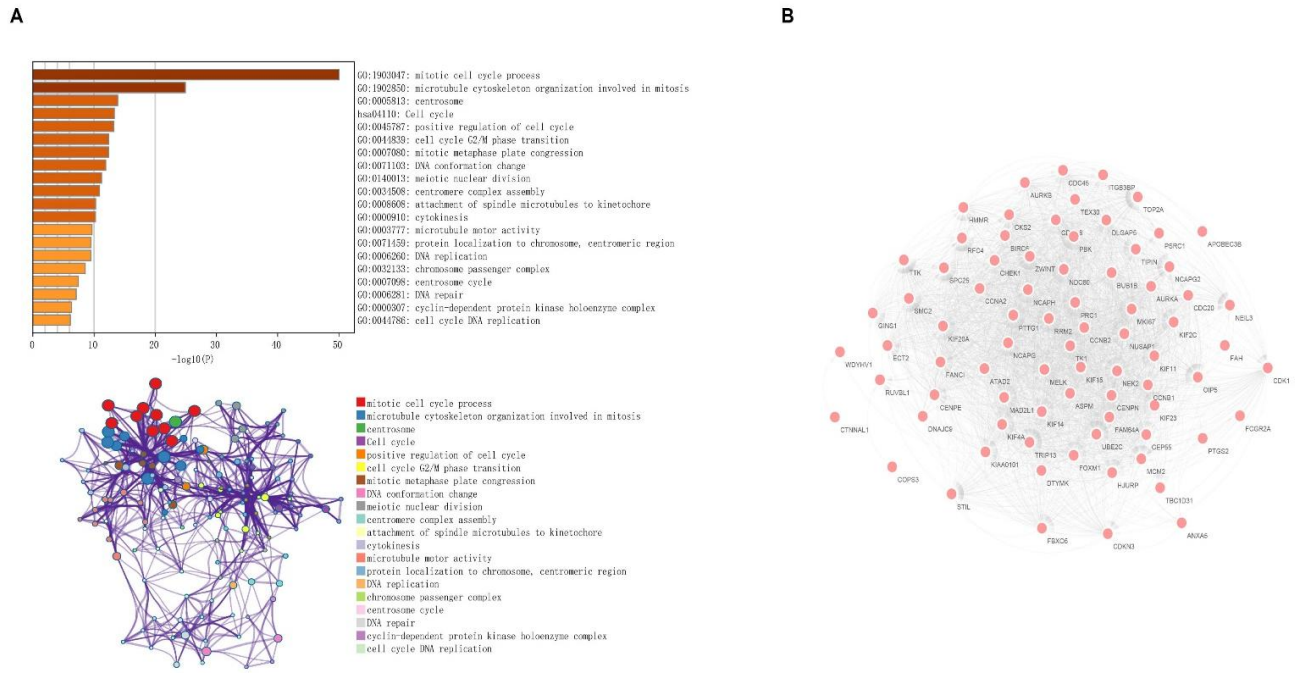


Figure 4. (A) the top half of Figure 4A: enrichment analysis by metascape database showed that genes were mainly enriched in mitotic cell cycle process, microtubule cytoskeleton organization involved in mitosis, centrosome, cell cycle, positive regulation of cell cycle, cell cycle G2/M phase transition, mitotic metaphase plate congression, DNA conformation change, meiotic nuclear division and other pathways; the bottom half of Figure 4A: cluster network of enriched pathways, where nodes sharing the same cluster are usually close to each other; (B) the interaction network which was obtained from the string database to obtain the interactions between proteins and then visualized by cytoscape.

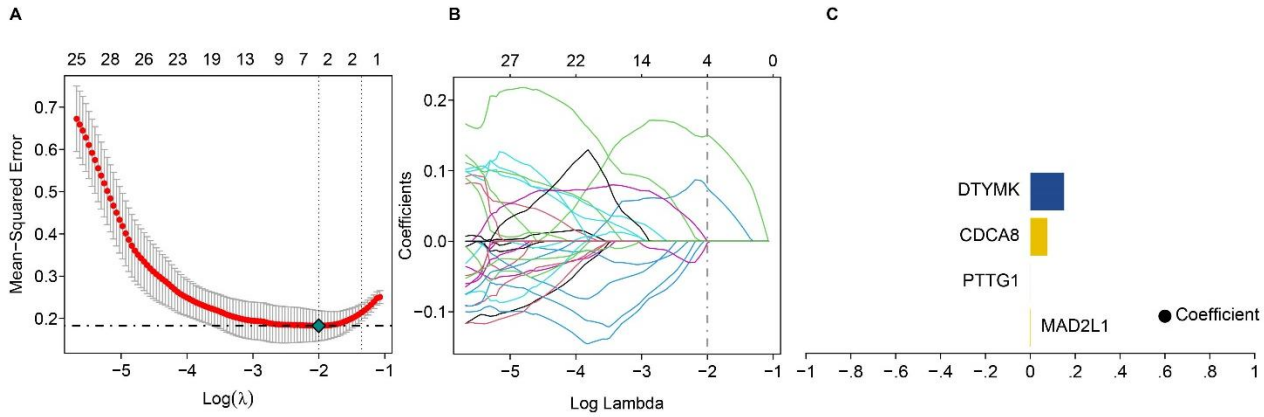


Figure 5. Screening of SCC core genes. (A) Ten-fold cross-validation of tuning parameter selection in the LASSO model. (B) LASSO coefficient distribution of differential genes. (C) Coefficients of the four Lasso (core) genes.

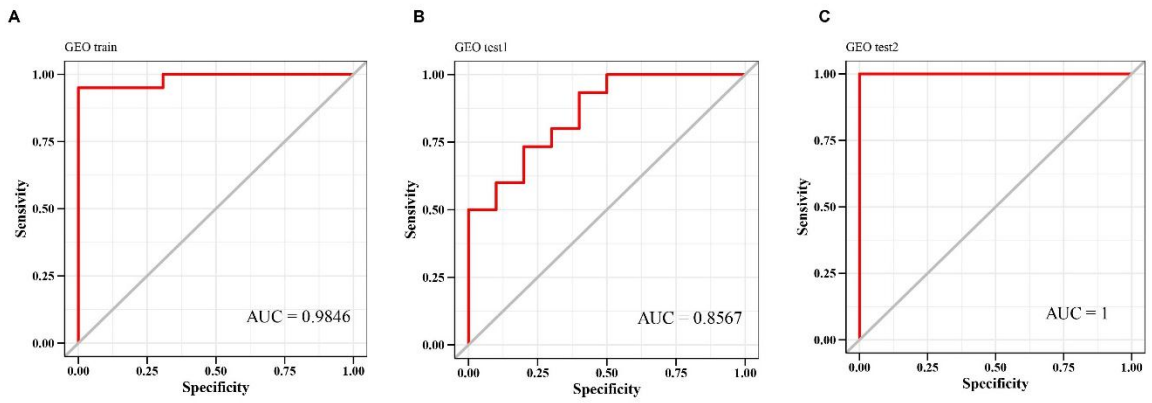


Figure 6. (A) constructing ROC curves of the prediction model using the training set (GSE42677 and GSE45164), area under the AUC curve is 0.9846; (B) further validating the ROC curves of the diagnosis model using the validation set GSE45216, area under the AUC curve is 0.8567; (C) further validating the ROC curves of the diagnosis model using the validation set GSE53462, area under the AUC curve is 1.

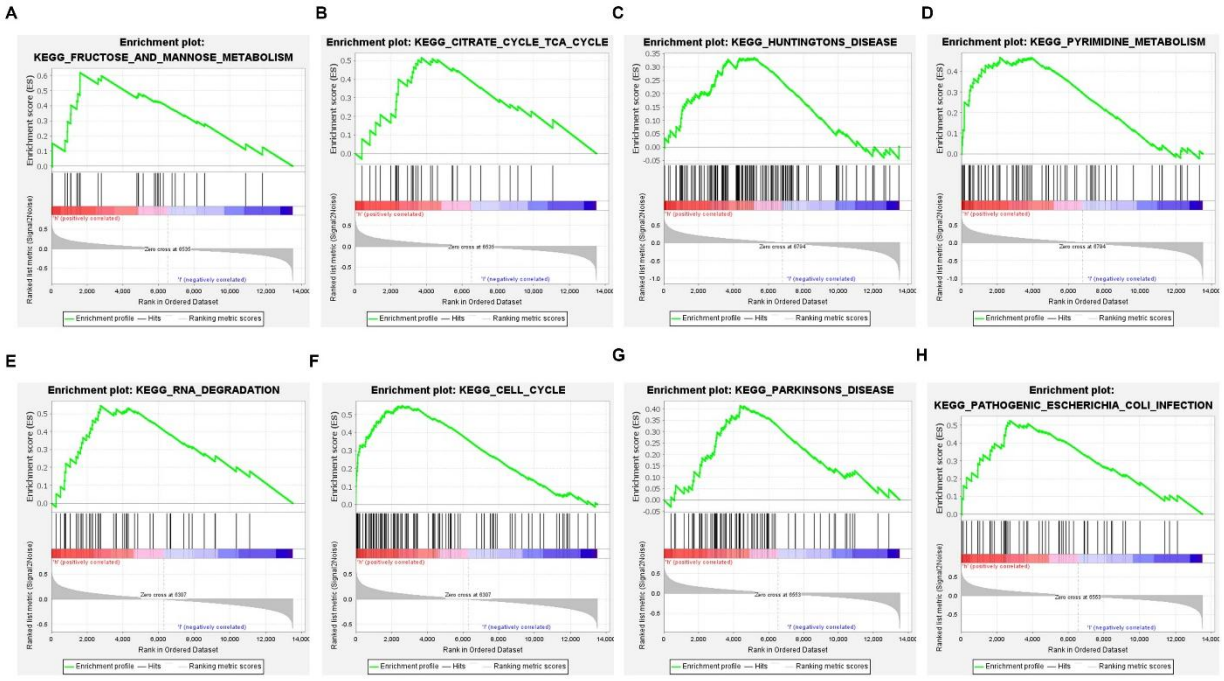
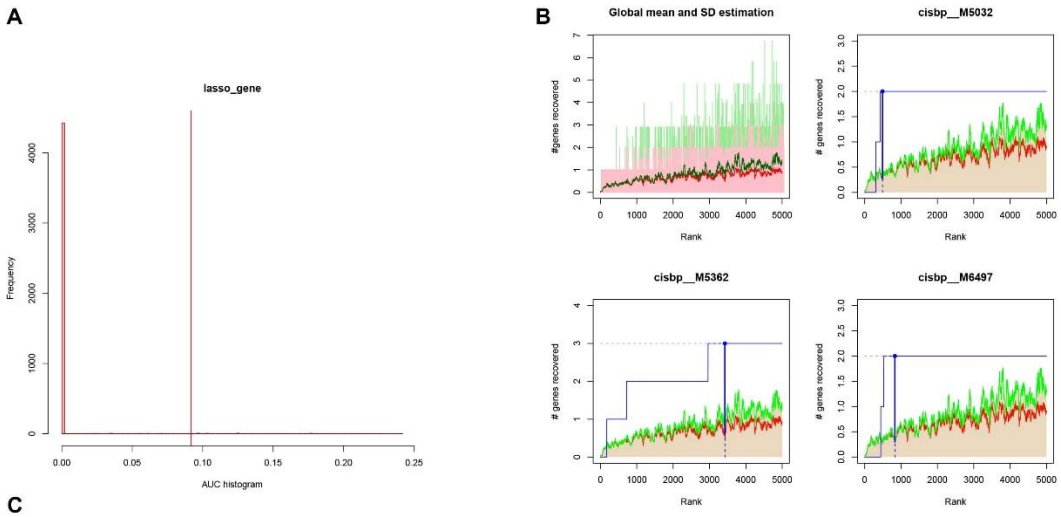


Figure 7. Gene set enrichment analysis of four core genes. (A) highly expressed CDCA8-enriched pathway (KEGG FRUCTOSE AND MANNOSE METABOLISM) (B) highly expressed CDCA8-enriched pathway (KEGG CITRATE CYCLE TCA CYCLE); (C) highly expressed DTYMK-enriched pathway (KEGG HUNTINGTONS DISEASE); (D) pathway highly enriched in DTYMK (KEGG PYRIMIDINE METABOLISM); (E) pathway highly enriched in MAD2L1 (KEGG RNA DEGRADATION); (F) pathway highly enriched in MAD2L1 (KEGG CELL CYCLE); (G) pathway enriched by high expression of PTTG1 (KEGG PARKINSONS DISEAS); (H) pathway enriched by high expression of PTTG1 (KEGG PATHOGENIC ESCHERICHIA COLI INFECTION).



C

	logo	geneSet	motif	NES	AUC	TF_highConf	nEnrGenes	rankAtMax
	All	All	All	All	All	All	All	All
1		lasso_gene	cisbp__M5032	7.9	0.334	RORA; RORB; RORC (inferredBy_Orthology).	2	491
2		lasso_gene	cisbp__M5362	7.03	0.299	E2F7 (directAnnotation).	3	3424
3		lasso_gene	cisbp__M6497	6.62	0.282	STAT6 (directAnnotation).	2	831
4		lasso_gene	cisbp__M1541	5.95	0.255		2	1383
5		lasso_gene	cisbp__M5791	5.94	0.254	RUNX3 (directAnnotation).	2	831

Figure 8. Analysis of Motif transcriptional regulation of core genes. (A,B) the red line is the mean of the recovery curve of each motif, the green line is the mean ± standard deviation (sd), and the blue line is the recovery curve of the current motif. The maximum distance point (mean ± sd) between the current motif and the green curve is the maximum enrichment level selected; (C) four core genes enriched to motifs. NES: normalized enrichment score of motifs in the gene set, AUC: area under the curve (used to calculate NES).

We did a differential study of immune regulatory genes and discovered that HLA-DMA, HLA-DOA, HLA-DOB, and HLA-DQB2 expression levels were significantly different between the two patient groups (Figure 10A). To investigate the relationship between essential genes and immune regulation, we used Pearson correlation analysis. The results indicated that PTTG1 and HLA-DMA were considerably favorably correlated (Pearson r = 0.62), whereas CDCA8 was strongly negatively connected with HLA-DQB2 (Pearson r = −0.63) (Figure 10B). We used the Human Protein Atlas website (https://www.proteinatlas.org/) to download immunohistochemical staining tissue images of the four core genes in normal skin tissue and tumor tissue, from which we can see the significant differential expression of the four core genes between normal and tumor tissues, further validating the accuracy of our model (Figure 11A–D).

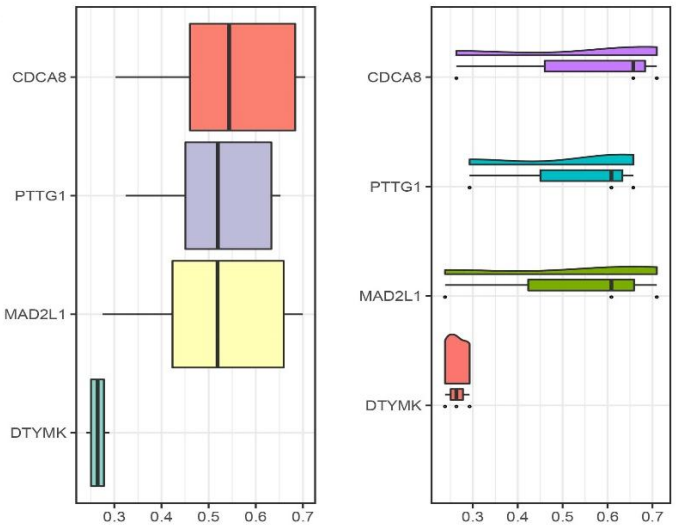


Figure 9. Screening of critical genes by GO semantic similarity showed that CDCA8 and PTTG1 are critical in the whole network. The upper and lower limits of the boxes show the 75th and 25th percentiles, and the line in the middle of the box indicates the mean value of similarity. The top two proteins are considered to be key proteins.

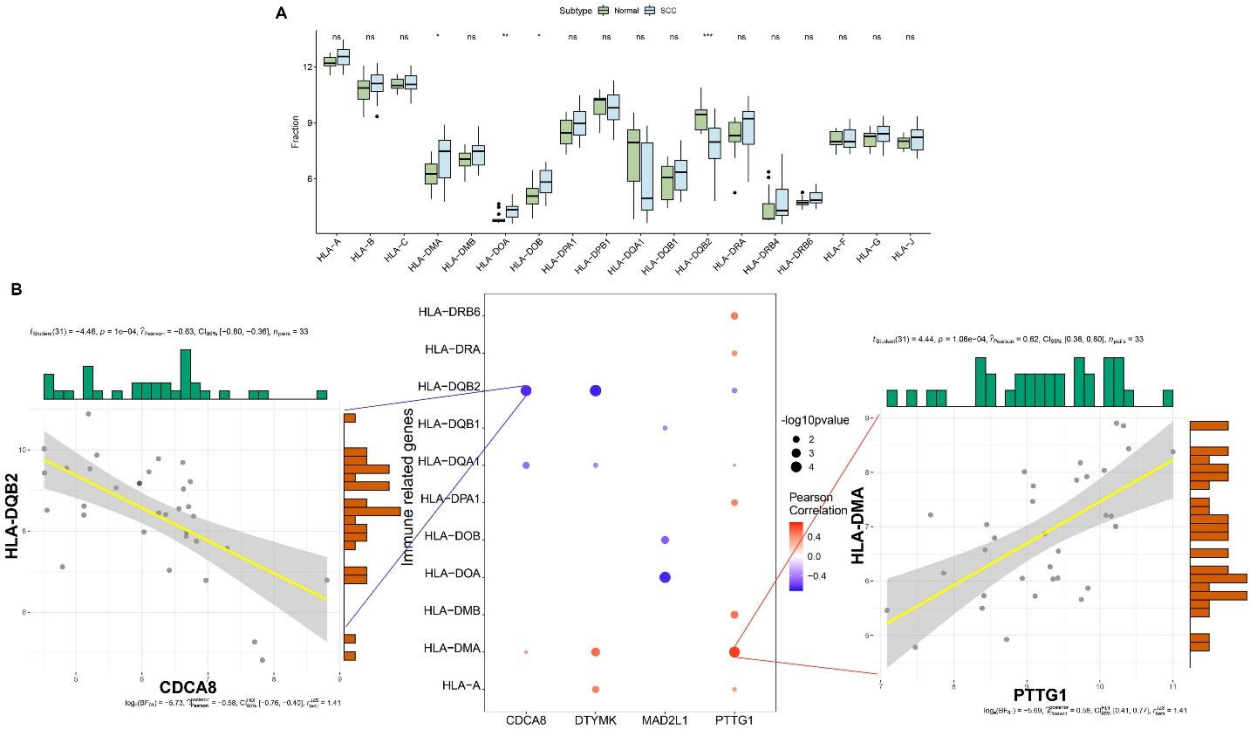


Figure 10. Correlation analysis of core genes and immune-related genes. (A) Differential analysis of immune regulatory genes showed that HLA-DMA, HLA-DOA, HLA-DOB and HLA-DQB2 were significantly different in the two groups of patients. * denotes $P < 0.05$, ** denotes $P < 0.01$, *** denotes $P < 0.001$; (B) Correlation analysis of critical genes PTTG1 and CDCA8 showed that PTTG1 and HLA-DMA were significantly positively correlated, and CDCA8 was significantly negatively correlated with HLA-DQB2.

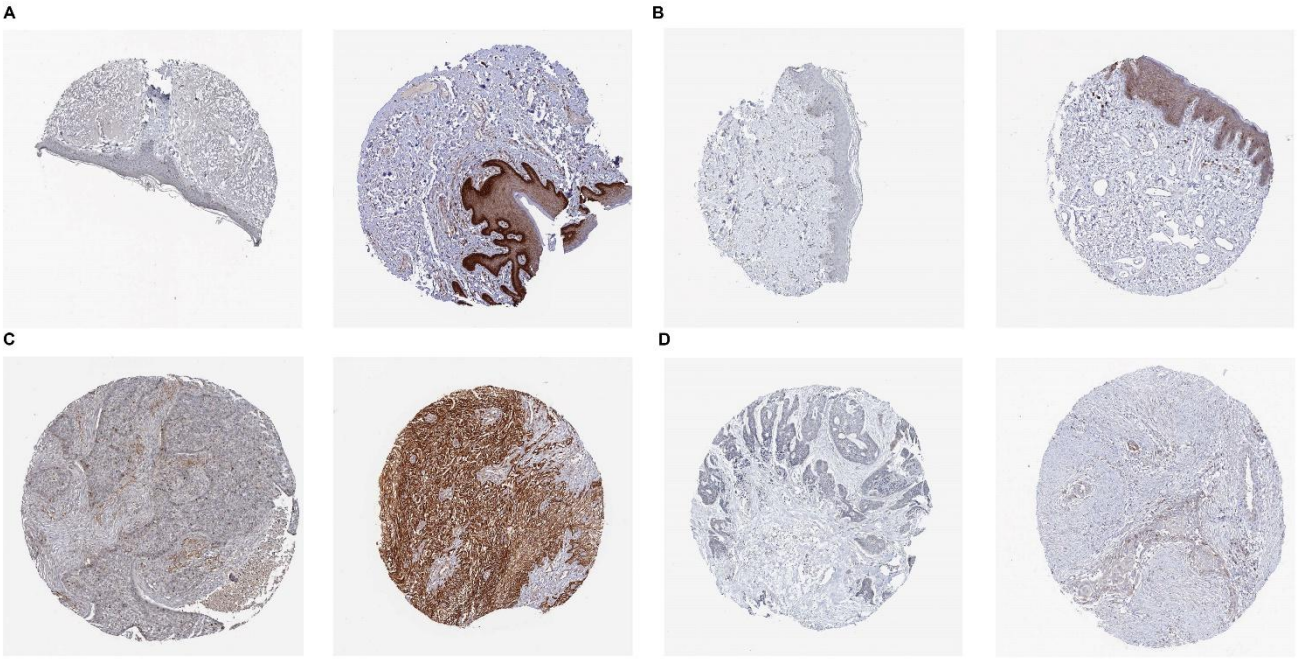


Figure 11. Immunohistochemical expression results of 4 core genes from Human Protein Atlas (https://www.proteinatlas.org/). (A) *CDCA8* expression in the normal and tumor tissue (HPA028058); (B) *DTYMK* expression in the normal and tumor tissue (HPA042593); (C) *MAD2L1* expression in in the normal and tumor tissue (HPA003348); (D) *PTTG1* expression in the normal and tumor tissue (HPA045034).

4. Discussion

Epithelial tumors account for 90% of human malignancies, and cSCC serves as the prototype of epithelial tumors due to its basic features of tissue polarity destruction and basement membrane infiltration. CSCC, the second most common cancer in the United States, affects more than one million people each year [32]. CSCC has a decent prognosis in its early stages, but about 4% of patients acquire lymph node metastases, and 1.5 percent of patients with advanced metastases die [3]. Despite the efficacy of chemotherapy, immunotherapy, and surgery in treating cSCC, its prognosis remains dismal at the moment [33]. Thus, it is critical to understand the tumor microenvironment and pathophysiology of cSCC to enhance the disease's prognosis and attain precision medicine.

We used the GEO dataset on cutaneous squamous cell cancer in this investigation. The study included 13 patients from the control group and 20 patients from the tumor group after batch effects between datasets were reduced using the Combat algorithm. We used the CIBERSORT method to measure immune cells in cSCC and built a WGCNA network based on patient T-cell correlations, and we discovered that the salmon module was the most significantly associated with T cells after evaluating nine gene modules and features. We developed a prediction model based on these four core genes shortly after using lasso regression to screen the modules of salmon for traits and discovered that DTYMK, CDCA8, PTTG1 and MAD2L1 play key roles in cSCC development and progression. The ROC curves of the training and validation sets showed good model stability. We

studied the specific signaling pathways enriched in the four core genes using GSEA analysis to infer illness progression biological processes. RORA, RORB and RORC were identified as significant transcription factors after enrichment analysis and Motif-TF annotation of transcription factors regulating the core genes. Following that, we ran a GO semantic similarity screen on them and discovered that CDCA8 and PTTG1 play important roles. We also looked at the association between core genes and immunomodulation and discovered that PTTG1 was positively correlated with HLA-DMA, whereas CDCA8 was dramatically negatively correlated with HLA-DQB2. Finally, we double-checked our findings using the HPA (Human Protein Atlas) online database.

We analyzed immune cell infiltration in the tumor microenvironment of cSCC. We discovered that there was a significant difference between the tumor and regular groups in B cells memory ($P = 0.016$), T cells regulatory (Tregs) ($P = 0.001$), Mast cells resting ($P = 0.002$), Mast cells activated ($P = 0.02$), and several other immune microenvironment factors. All of these immune cells play a relatively critical role in the development of cSCC. Most cancers, including cSCC, contain Treg T cells, and their presence usually indicates a bad prognosis [34,35]. Treg T cells can inhibit cytokine release and proliferation in effector cells. They can also suppress the activation of effectors by influencing antigen-presenting cells. Skin infiltration can be induced by UV irradiation, which causes infiltrating T cells to generate Treg T cells. TGF-β is produced by malignant cells during the carcinogenesis of cSCC and can cause infiltrating effector T cells to become Treg T cells [36,37]. Treg T cells were discovered in greater numbers in cSCC than in normal skin, and in moderately as well as poorly differentiated cSCC than in highly differentiated cSCC [36]. Macrophages are the most common type of leukocyte found in most cancers. Tumor-associated macrophages (TAMs) are altered by paracrine signals such as TGF-β and IFN-γ in late stages of cSCC carcinogenesis, resulting in inefficient phagocytosis and antigen presentation, and hence acquire an immunoregulatory phenotype [38,39]. Around and within the cut margins of surgically excised cSCCs, we discovered a substantial number of M1 and M2 macrophage phenotypes [40,41]. TAMs directly stimulate tumor growth by secreting pro-vascular growth factors like vascular endothelial growth factor (VEGF) and matrix metalloproteases (MMP), which breach the normal tissue barrier, allowing for additional invasion [40,42]. Because of the ability to create antibodies and effector cytokines, mediate antigen presentation, and contribute to T cell polarization, B cells perform an anti-tumor role [43,44]. Furthermore, B cells are involved in tumor growth and cancer immunity [45]. B cell depletion causes oncogenic failure in animal models of cSCC, owing to their synthesis of immunoglobulins with chemotactic and pro-inflammatory effects, as well as the deposition of stromal immune complexes early in tumor growth, leading to myeloid activation via Fc receptors [46,47]. TNF-α and IL-10 can also be used by B cells to inhibit the immunological response to cSCC [48]. Furthermore, it has been discovered that during chemical carcinogenesis, B cells can block the development of cSCC by generating IgE [49].

We discovered that DTYMK, CDCA8, PTTG1 and MAD2L1 play critical roles in cSCC development and progression. DTYMK (deoxythymidylate kinas) may be used to predict the prognosis of non-small cell lung cancer and hepatocellular carcinoma patients. It is closely associated with cell cycle and acid metabolism pathways and suppresses the immune microenvironment during tumorigenesis. Additionally, overexpression of DTYMK in some patients increases their sensitivity to chemotherapeutic agents, suggesting that it may be a potential therapeutic target [50,51]. CDCA8 (cell division cycle associated 8) may provide a novel molecular method for the initial treatment of hepatocellular carcinoma and preventing metastases. CDCA8

silencing reduces the growth of hepatocellular carcinoma via restoring the ATF3 oncogene and inactivating oncogenic AKT/β-linked protein signaling, whereas CDCA8 overexpression promotes tumor cell proliferation. Additionally, CDCA8 is a critical regulator of mitosis and is being investigated as a possible predictive biomarker for many malignancies, including lung, breast, and colon [52]. PTTG1 (pituitary tumor-transforming gene-1) is highly expressed in esophageal, laryngeal, lung, and hepatocellular carcinomas [53,54], and is associated with the degree of differentiation, invasion, and metastasis of the tumor [55,56]. PTTG1 is a chromosomal stability regulator that can aggravate tumor development progression, and inhibiting PTTG1 decreases cancer cell proliferation and promotes apoptosis. Additionally, PTTG1 expression is increased in hematologic malignant proliferative disorders such as multiple myeloma and is significantly related to a bad prognosis [57]. MAD2L1 (mitotic arrest deficient 2 like 1) is an emerging target for anticancer therapy because it plays a critical role in mitosis, the cell cycle checkpoint, and the DNA damage response [58]. MAD2L1 has been demonstrated to be a novel prognostic factor for lung adenocarcinoma and squamous cell carcinoma, and it also interacts with CDC20 and BUB1B [59–61]. MAD2L1 and MAD2L2 can result in aberrant salivary ductal cancer formation [62,63].

In summary, our current work elucidates the immune cell types and interactions in the tumor microenvironment of cSCC, identifies critical components, and identifies enrichment routes via core genes. Meanwhile, we conducted differential analysis on immune regulatory genes and connected them with disease-associated genes.

5. Conclusions

In summary, we discovered that DTYMK, CDCA8, PTTG1 and MAD2L1 are the critical genes in cSCC, with CDCA8 and PTTG1 being particularly important. The primary regulators are RORA, RORB, and RORC. B cells memory, T cells regulatory, Mast cells resting, and Mast cells activation were significantly different between the tumor and regular groups. We compared essential genes with immune regulatory genes and discovered a strong positive correlation between PTTG1 and HLA-DMA and a significant negative correlation between CDCA8 and HLA-DQB2.

Conflicts of interest

All authors declare no conflicts of interest in this paper.

References

1. C. Fitzmaurice, D. Abate, N. Abbasi, H. Abbastabar, F. Abd-Allah, O. Abdel-Rahman, et al., Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: a Systematic analysis for the global burden of disease study, *JAMA Oncol.*, **5** (2019), 1749–1768. https://doi.org/10.1001/jamaoncol.2019.2996
2. H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J. Clin.*, **71** (2021), 209–249. https://doi.org/10.3322/caac.21660

3. P. S. Karia, J. Han, C. D. Schmults, Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012, *J. Am. Acad. Dermatol.*, **68** (2013), 957–966. https://doi.org/10.1016/j.jaad.2012.11.037
4. J. M. Janus, R. F. L. O'Shaughnessy, C. Harwood, T. Maffucci, Phosphoinositide 3-Kinase-Dependent signalling pathways in cutaneous squamous cell carcinomas, *Cancers*, **9** (2017), 86. https://doi.org/10.3390/cancers9070086
5. M. Piipponen, R. Riihilä, L. Nissinen, V. Kähäri, The role of p53 in progression of cutaneous squamous cell carcinoma, *Cancers*, **13** (2021), 4507. https://doi.org/10.3390/cancers13184507
6. A. Boutros, F. Cecchi, E. Tanda, E. Croce, R. Gili1, L. Arecco, et al., Immunotherapy for the treatment of cutaneous squamous cell carcinoma, *Front. Oncol.*, **11** (2021), 733917. https://doi.org/10.3389/fonc.2021.733917
7. Y. Sawada, M. Nakamura, Daily lifestyle and cutaneous malignancies, *Int. J. Mol. Sci.*, **22** (2021), 5227. https://doi.org/10.3390/ijms22105227
8. K, Suozzi, J. Turban, M. Girardi, Cutaneous photoprotection: a review of the current status and evolving strategies, *Yale J. Biol. Med.*, **93** (2020), 55–67.
9. C. Flower, D. Gaskin, S. Bhamjee, Z. Bynoe, High-risk variants of cutaneous squamous cell carcinoma in patients with discoid lupus erythematosus: a case series, *Lupus*, **22** (2013), 736–739. https://doi.org/10.1177%2F0961203313490243
10. K. K. Das, A. Chakaraborty, A. Rahman, S. Khandkar, Incidences of malignancy in chronic burn scar ulcers: experience from Bangladesh, *Burns*, **41** (2015), 1315–1321. https://doi.org/10.1016/j.burns.2015.02.008
11. T. J. Knackstedt, L. K. Collins, Z. Li, S. Yan, F. Samie, Squamous cell carcinoma arising in hypertrophic lichen planus: a review and analysis of 38 cases, *Dermatol. Surg.*, **41** (2015), 1411–1418. http://doi.org/10.1097/DSS.0000000000000565
12. J. Xing, Z. Jia, Y. Xu, M. Chen, Z. Yang, Y. Chen, et al., KLF9 (Kruppel Like Factor 9) induced PFKFB3 (6-Phosphofructo-2-Kinase/Fructose-2, 6-Biphosphatase 3) downregulation inhibits the proliferation, metastasis and aerobic glycolysis of cutaneous squamous cell carcinoma cells, *Bioengineered*, **12** (2021), 7563–7576. https://doi.org/10.1080/21655979.2021.1980644
13. J. G. Newman, M. A. Hall, S. J. Kurley, R. Cook, A. S. Farberg, J. L. Geiger, et al., Adjuvant therapy for high-risk cutaneous squamous cell carcinoma: 10-year review, *Head Neck*, **43** (2021), 2822–2843. https://doi.org/10.1002/hed.26767
14. J. Pang, H. Pan, C. Yang, P. Meng, W. Xie, J. Li, et al., Prognostic value of immune-related multi-incRNA signatures associated with tumor microenvironment in esophageal cancer, *Front. Genet.*, **12** (2021), 722601. https://dx.doi.org/10.3389%2Ffgene.2021.722601
15. Y. Pan, H. Han, K. E. Labbe, H. Zhang, W. Wong, Recent advances in preclinical models for lung squamous cell carcinoma, *Oncogene*, **40** (2021), 2817–2829. https://doi.org/10.1038/s41388-021-01723-7
16. A. Elmusrati, J. Wang, C. Y. Wang, Tumor microenvironment and immune evasion in head and neck squamous cell carcinoma, *Int. J. Oral. Sci.*, **13** (2021), 24. https://doi.org/10.1038/s41368-021-00131-7
17. T. Suwa, M. Kobayashi, J. M. Nam, H, Harada, Tumor microenvironment and radioresistance, *Exp. Mol. Med.*, **53** (2021), 1029–1035. https://doi.org/10.1038/s12276-021-00640-9
18. S. Paget, The distribution of secondary growths in cancer of the breast, *Cancer Metastasis Rev.*, **8** (1889), 98–101.

19. H. Wang, M. M. H. Yung, H. Y. S. Ngan, K. Chan, D. W. Chan, The impact of the tumor microenvironment on macrophage polarization in cancer metastatic progression, *Int. J. Mol. Sci.*, **22** (2021), 6560. https://doi.org/10.3390/ijms22126560
20. J. Zhuyan, M. Chen, T. Zhu, X. Bao, T. Zhen, K. Xing, et al., Critical steps to tumor metastasis: alterations of tumor microenvironment and extracellular matrix in the formation of pre-metastatic and metastatic niche, *Cell Biosci.*, **10** (2020), 89. https://doi.org/10.1186/s13578-020-00453-9
21. Y. Xie, F. Xie, L. Zhang, X. Zhou, J. Huang, F. Wang, et al., Targeted anti-tumor immunotherapy using tumor infiltrating cells, *Adv. Sci.*, e2101672. https://doi.org/10.1002/advs.202101672
22. M. Akhtar, A. Haider, S. Rashid, A. Ai-Nabet, Paget's "Seed and Soil" theory of cancer metastasis: an idea whose time has come, *Adv. Anat. Pathol.*, **26** (2019), 69–74. https://doi.org/10.1097/PAP.0000000000000219
23. G. Yan, L. Li, S. Zhu, Y. Wu, Y. Zhu, L. Zhu, et al., Single-cell transcriptomic analysis reveals the critical molecular pattern of UV-induced cutaneous squamous cell carcinoma, *Cell Death Dis.*, **13** (2022), 23. https://doi.org/10.1038/s41419-021-04477-y
24. A. Ji, A. Rubin, K. Thrane, S. Jiang, D. L. Reynolds, R. M. Meyers, et al., Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma, *Cell*, **182** (2020), 497–514. https://doi.org/10.1016/j.cell.2020.05.039
25. C. B. Steen, C. L. Liu, A. A. Alizadeh, A. M. Newman, Profiling cell type abundance and expression in bulk tissues with CIBERSORTx, *Methods Mol. Biol.*, **2117** (2020), 135–157. https://doi.org/10.1007/978-1-0716-0301-7_7
26. J. L. Sevilla, V. Segura, A. Podhorski, E. Guruceaga, J. M. Mato, L. A. Martinez-Cruz, et al., (2005) Correlation between gene expression and GO semantic similarity, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, **2** (2005), 330–338. https://doi.org/10.1109/TCBB.2005.50
27. S. Jain, G. D. Bader, An improved method for scoring protein-protein interactions using semantic similarity within the gene ontology, *BMC Bioinf.*, **11** (2010), 562. https://doi.org/10.1186/1471-2105-11-562
28. X. Guo, C. D. Shriver, H. Hu, M. N. Liebman, Analysis of metabolic and regulatory pathways through Gene Ontology-derived semantic similarity measures, in *AMIA Annual Symposium Proceedings, American Medical Informatics Association,* (2005), 972.
29. P. M. Tedder, J. R. Bradford, C. J. Needham, G. A. McConkey, A. J. Bulpitt, D. R. Westhead, Gene function prediction using semantic similarity clustering and enrichment analysis in the malaria parasite Plasmodium falciparum, *Bioinformatics*, **26** (2010), 2431–2437. https://doi.org/10.1093/bioinformatics/btq450
30. G. Yu, F. Li, Y. Qin, X. Bo, Y. Wu, S. Wang, GOSemSim: an R package for measuring semantic similarity among GO terms and gene products, *Bioinformatics*, **26** (2010), 976–978. https://doi.org/10.1093/bioinformatics/btq064
31. J. Z. Wang, Z. Du, R. Payattakool, P. S. Yu, C. F. Chen, A new method to measure the semantic similarity of GO terms, *Bioinformatics*, **23** (2007), 1274–1281. https://doi.org/10.1093/bioinformatics/btm087
32. E. Rognoni, M. Widmaier, M. Jakobson, R. Ruppert, S. Ussar, D. Katsougkri, Kindlin-1 controls Wnt and TGF-β availability to regulate cutaneous stem cell proliferation, *Nat. Med.*, **20** (2014), 350–359. https://doi.org/10.1038/nm.3490

33. M. Lai, R. Pampena, L. Cornacchia, G. Odorici, A. Piccerillo, G. Pellacani, et al., Cutaneous squamous cell carcinoma in patients with chronic lymphocytic leukemia: a systematic review of the literature, *Int. J. Dermatol*., **2021** (2021). https://doi.org/10.1111/ijd.15813
34. H. B. Jie, P. J. Schuler, S. C. Lee, R. M. Srivastava, A. Argiris, S. Ferrone, et al., CTLA-4$^+$ regulatory T cells increased in cetuximab-treated head and neck cancer patients suppress NK cell cytotoxicity and correlate with poor prognosis, *Cancer Res.*, **75** (2015), 2200–2210. https://doi.org/10.1158/0008-5472.CAN-14-2788
35. S. Z. Lin, K. J. Chen, Z. Y. Xu, H. Chen, L. Zhou, H. Y. Xie, et al., Prediction of recurrence and survival in hepatocellular carcinoma based on two Cox models mainly determined by FoxP3+ regulatory T cells, *Cancer Prev. Res.*, **6** (2013), 594–602. https://doi.org/10.1158/1940-6207.CAPR-12-0379
36. B. Azzimonti, E. Zavattaro, M. Provasi, M. Vidali, A. Conca, E. Catalano, et al., Intense Foxp3+ CD25+ regulatory T-cell infiltration is associated with high-grade cutaneous squamous cell carcinoma and counterbalanced by CD8+/Foxp3+ CD25+ ratio, *Br. J. Dermatol.*, **172** (2014), 64–73. https://doi.org/10.1111/bjd.13172
37. S. M. Gorsch, V. A. Memoli, T. A. Stukel, L. I. Gold, B. A. Arrick, Immunohistochemical staining for transforming growth factor beta 1 associates with disease progression in human breast cancer, *Cancer Res.*, **52** (1992), 6949–6952.
38. M. Ponzoni, F. Pastorino, D. Di Paolo, P. Perri, C. Brignole, Targeting macrophages as a potential therapeutic intervention: impact on inflammatory diseases and cancer, *Int. J. Mol. Sci.*, **19** (2018), 1953. https://doi.org/10.3390/ijms19071953
39. L. Nissinen, M. Farshchian, P. Riihilä, V. Kähäre, New perspectives on role of tumor microenvironment in progression of cutaneous squamous cell carcinoma, *Cell Tissue Res.*, **365** (2016), 691–702. https://doi.org/10.1007/s00441-016-2457-z
40. J. S. Pettersen, J. Fuentes-Duculan, M. Suárez-Fariñas, K. C. Pierson, A. Pitts-Kiefer, L. Fan, et al., Tumor-associated macrophages in the cutaneous SCC microenvironment are heterogeneously activated, *J. Invest. Dermatol.*, **131** (2011), 1322–1330. https://doi.org/10.1038/jid.2011.9
41. M. Takahara, S. Chen, M. Kido, S. Takeuchi, H. Uchi, Y. Tu, et al., Stromal CD10 expression, as well as increased dermal macrophages and decreased Langerhans cells, are associated with malignant transformation of keratinocytes, *J. Cutan. Pathol.*, **36** (2009), 668–674. https://doi.org/10.1111/j.1600-0560.2008.01139.x
42. D. Moussai, H. Mitsui, J. S. Pettersen, K. C. Pierson, K. R. Shah, M. Suárez- Fariñas, et al., The human cutaneous squamous cell carcinoma microenvironment is characterized by increased lymphatic density and enhanced expression of macrophage-derived VEGF-C, *J. Invest. Dermatol.*, **131** (2011), 229–236. https://doi.org/10.1038/jid.2010.266
43. C. A. Janeway, J. Ron, M. E. Katz, The B cell is the initiating antigen-presenting cell in peripheral lymph nodes, *J. Immunol.*, **138** (1987), 1051–1055.
44. D. P. Harris, L. Haynes, P. C. Sayles, D. K. Duso, S. M. Eaton, N. M. Lepak, et al., Reciprocal regulation of polarized cytokine production by effector B and T cells, *Nat. Immunol.*, **1** (2000), 475–482. https://doi.org/10.1038/82717
45. A. Sarvaria, J. A. Madrigal, A. Saudemont, B cell regulation in cancer and anti-tumor immunity, *Cell Mol. Immunol.*, **14** (2017), 662–674. https://doi.org/10.1038/cmi.2017.35

46. P. Andreu, M. Johansson, N. Affara, F. Pucci, T. Tan, S. Junankar, et al., FcRgamma activation regulates inflammation-associated squamous carcinogenesis, *Cancer Cell*, **17** (2010), 121–134. https://doi.org/10.1016/j.ccr.2009.12.019
47. K. W. de Visser, L. V. Korets, L. M. Coussens, De novo carcinogenesis promoted by chronic inflammation is B lymphocyte dependent, *Cancer Cell*, **7** (2005), 411–423. https://doi.org/10.1016/j.ccr.2005.04.014
48. T. Schioppa, R. Moore, R. G. Thompson, F. R. Balkwill, B regulatory cells and the tumor-promoting actions of TNF-α during squamous carcinogenesis, *Proc. Natl. Acad. Sci.*, **108** (2011), 10662–10667. https://doi.org/10.1073/pnas.1100994108
49. G. Crawford, M. D. Hayes, R. C. Seoane, S. Ward, T. Dalessandri, C. Lai, et al., Epithelial damage and tissue γδ T cells promote a unique tumor-protective IgE response, *Nat. Immunol.*, **19** (2018), 859–870. https://doi.org/10.1038/s41590-018-0161-8
50. T. Zhou, R. Qin, S. Shi, H. Zhang, C. Niu, G. Ju, et al., DTYMK promote hepatocellular carcinoma proliferation by regulating cell cycle, *Cell Cycle*, **20** (2021), 1681–1691. https://doi.org/10.1080/15384101.2021.1958502
51. Y. Guo, W. Luo, S. Huang, W. Zhao, H. Chen, Y. Ma, et al., DTYMK expression predicts prognosis and chemotherapeutic response and correlates with immune infiltration in hepatocellular carcinoma, *J. Hepatocell Carcinoma*, **8** (2021), 871–885. https://dx.doi.org/10.2147%2FJHC.S312604
52. T. Jeon, M. J. Ko, Y. R. Seo, S. J. Jung, D. Seo, S. Y. Park, et al., Silencing CDCA8 suppresses hepatocellular carcinoma growth and stemness via restoration of ATF3 tumor suppressor and inactivation of AKT/β-catenin signaling, *Cancers*, **13** (2021), 1055. https://doi.org/10.3390/cancers13051055
53. G. Vlotides, T. Eigler, S. Melmed, Pituitary tumor-transforming gene: physiology and implications for tumorigenesis, *Endocr. Rev.*, **28** (2007), 165–186. https://doi.org/10.1210/er.2006-0042
54. H. Hong, Z. Jin, T. Qian, X. Xu, X. Zhu, Q. Fei, et al., Falcarindiol enhances cisplatin chemosensitivity of hepatocellular carcinoma via down-regulating the STAT3-modulated PTTG1 pathway, *Front. Pharmacol.*, **12** (2021), 656697. https://dx.doi.org/10.3389%2Ffphar.2021.656697
55. S. W. Chen, H. F. Zhou, H. J. Zhang, R. He, Z. Huang, Y. Dang, et al., The clinical significance and potential molecular mechanism of PTTG1 in esophageal squamous cell carcinoma, *Front. Genet.*, **11** (2021), 583085. https://doi.org/10.3389/fgene.2020.583085
56. Z. Chen, K. Cao, Y. Hou, F. Lu, L. Li, L. Wang, et al., PTTG1 knockdown enhances radiation-induced antitumour immunity in lung adenocarcinoma, *Life Sci.*, **277** (2021), 119594. https://doi.org/10.1016/j.lfs.2021.119594
57. J. E. Noll, K. Vandyke, D. R. Hewett, K. M. Mrozik, R. J. Bala, S. A. Williams, et al., PTTG1 expression is associated with hyperproliferative disease and poor prognosis in multiple myeloma, *J. Hematol. Oncol.*, **8** (2015), 106. https://doi.org/10.1186/s13045-015-0209-2
58. R. Wei, Z. Wang, Y. Zhang, B. Wang, N. Shen, E. Li, et al., Bioinformatic analysis revealing mitotic spindle assembly regulated NDC80 and MAD2L1 as prognostic biomarkers in non-small cell lung cancer development, *BMC Med. Genomics*, **13** (2020), 112. https://doi.org/10.1186/s12920-020-00762-5

59. M. Vleugel, T. A. Hoek, E. Tromer, T. Sliedrecht, V. Groenewold, M. Omerzu, et al., Dissecting the roles of human BUB1 in the spindle assembly checkpoint, *J. Cell Sci.*, **128** (2015), 2975–2982. https://doi.org/10.1242/jcs.169821
60. Y. H. Ko, J. H. Roh, Y. I. Son, M. K. Chung, J. Y. Jang, H. Byun, et al., Expression of mitotic checkpoint proteins BUB1B and MAD2L1 in salivary duct carcinomas, *J. Oral Pathol. Med.*, **39** (2010), 349–355. https://doi.org/10.1111/j.1600-0714.2009.00835.x
61. M. Abal, A. Obrador-Hevia, K. P. Janssen, L. Casadome, M. Menendez, S. Carpentier, et al., APC inactivation associates with abnormal mitosis completion and concomitant BUB1B/MAD2L1 up-regulation, *Gastroenterology*, **132** (2007), 2448–2458. https://doi.org/10.1053/j.gastro.2007.03.027
62. Y. Wang, Z. Zhou, L. Chen, Y. Li, Z. Zhou, X. Chu, Identification of key genes and biological pathways in lung adenocarcinoma via bioinformatics analysis, *Mol. Cell Biochem.*, **476** (2021), 931–939. https://doi.org/10.1007/s11010-020-03959-5
63. R. Marima, R. Hull, C. Penny, Z. Dlamini, Mitotic syndicates Aurora Kinase B (AURKB) and mitotic arrest deficient 2 like 2 (MAD2L2) in cohorts of DNA damage response (DDR) and tumorigenesis, *Mutat. Res. Rev. Mutat. Res.*, **787** (2021), 108376. https://doi.org/10.1016/j.mrrev.2021.108376

AIMS Press