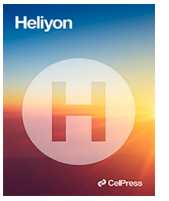
Heliyon 9 (2023) e19853

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Development and validation of a differentiation-related signature based on single-cell RNA sequencing data of immune cells in spinal cord injury

Jun Shang [a,b,1], Chao Ma [a,1], Han Ding [a,1], Guangjin Gu [a,1], Jianping Zhang [a], Min Wang [c], Ke Fang [a], Zhijian Wei [a,**], Shiqing Feng [a,b,*]

[a] International Science and Technology Cooperation Base of Spinal Cord Injury, Tianjin Key Laboratory of Spine and Spinal Cord Injury, Department of Orthopedics, Tianjin Medical University General Hospital, Tianjin, China
[b] Department of Orthopaedics, The Second Hospital of Shandong University, Shandong University, Jinan, Shandong, China
[c] Tianjin Key Laboratory of Lung Cancer Metastasis and the Tumor Microenvironment, Tianjin Lung Cancer Institute, Tianjin Medical University General Hospital, Tianjin, China

ARTICLE INFO

Keywords:
Spinal cord injury
Single cell trajectory analysis
Immune infiltration
Molecular subtypes

ABSTRACT

Background: After spinal cord injury (SCI), the native immune surveillance function of the central nervous system is activated, resulting in a substantial infiltration of immune cells into the affected tissue. While numerous studies have explored the transcriptome data following SCI and revealed certain diagnostic biomarkers, there remains a paucity of research pertaining the identification of immune subtypes and molecular markers related to the immune system post-spinal cord injury using single-cell sequencing data of immune cells.

Methods: The researchers conducted an analysis of spinal cord samples obtained at three time points (3,10, and 21 days) following SCI using the GSE159638 dataset. The SCI subsets were delineated through pseudo-time analysis, and differentiation related genes were identified after principal component analysis (PCA), cell clustering, and annotation techniques. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were employed to assess the differentiation-related genes (DRGs) across different subsets. The molecular subtypes of SCI were determined using consensus clustering analysis. To further explore and validate the correlation between the molecular subtypes and the immune microenvironment, the CIBERSORT algorithm was employed. High-value diagnostic gene markers were identified using LASSO regression, and their diagnostic sensitivity was assessed using receiver operating characteristic curves (ROC) and quantitative real-time polymerase chain reaction (qRT-PCR).

Results: Three SCI subsets were obtained, and differentiation-related genes were characterized. Within these subsets, two distinct molecular subtypes, namely C1 and C2, were identified. These subtypes demonstrated significant variations in terms of immune cell infiltration levels and the expression of immune checkpoint genes. Through further analysis, three candidate biomarkers (C1qa, Lgals3 and Cd63) were identified and subsequently validated.

* Corresponding author.Department of Orthopedics, Tianjin Medical University General Hospital, 154 Anshan Road, Heping District, Tianjin, 300052, China.
** Corresponding author.
E-mail addresses: weizhijian2002@126.com (Z. Wei), sqfeng@tmu.edu.cn (S. Feng).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2023.e19853
Received 11 April 2023; Received in revised form 2 August 2023; Accepted 4 September 2023
Available online 9 September 2023

Conclusions: Our study revealed a diverse immune microenvironment in SCI samples, highlighting the potential significance of C1qa, Lgals3 and Cd63 as immune biomarkers for diagnosing SCI. Moreover, the identification of immune checkpoints corresponding to the two molecular subtypes suggests their potential as targets for immunotherapy to enhance SCI repair in future interventions.

1. Introduction

Spinal cord injury (SCI), as a severe central neurological disorder, is characterized by the loss of sensorimotor function below the level of injury, limited recovery capacity, and various complications [1–3]. Recent epidemiological studies have reported a global incidence estimate of 27.04 million (24.98–30.15 million) cases of SCI [4]. This spinal cord injury results in paralysis, significantly diminishes quality of life, and imposes substantial economic burdens on both affected families of patients and society as a whole [5,6].

Post-SCI inflammation is complex process involving multiple cell types and inflammatory cytokines, such as tumor necrosis factor alpha (TNFα), interleukin-1β (IL-1β), and interleukin-6 (IL-6). While inflammation following SCI have some beneficial effects, the excessive infiltration of immune cells is primarily responsible for neurodegeneration [1,7,8]. The release of inflammatory cytokines and chemokines triggers an extensive infiltration of immune cells, including neutrophils, microglia, and macrophages of peripheral origin, which further produce additional inflammatory mediators. Although the extent of neuroinflammation is influenced by the severity of the initial injury, often resulting in an overreaction of the inflammatory process after spinal cord injury, ultimately leading to additional cell death [9,10]. Early decompression and high-dose methylprednisolone (steroid) are commonly used in clinical practice for treatment [11]. However, numerous reports show that the risks associated with steroid use outweigh the benefits [12]. In light of these concerns, researchers have been exploring alternative therapeutic approaches aimed at promoting neuronal differentiation from neural stem cells in situ, enhancing axonal elongation in residual neurons, and modulating the local inhibitory microenvironment [13–17]. Although some of these novel strategies show promise, none have yet proven clinical efficacy.

Numerous clinical studies have heighted the existence of varying degrees of recovery among patients with similar pathological conditions, underscoring the importance of patient heterogeneity and the need for precision therapy [18–21]. Given the local intricate and diverse immune microenvironment observed after spinal cord injury, it is essential to develop differentiation markers for immune cells post spinal cord injury in order to deepen our understanding of immune infiltration and the underlying molecular mechanisms. Such knowledge will promote the precise application of immunotherapy after spinal cord injury, ultimately improving both short-term and long-term prognosis in individuals with central nervous system (CNS) injuries.

To reveal the intricate functions and mechanisms of cells, the application of single-cell sequencing analysis has become increasingly prevalent as a cutting-edge technology [22,23]. By examining the trajectories of cell differentiation in single-cell RNA (scRNA) sequences, we successfully identify distinct cell subtypes within the immune microenvironment of spinal cord injury (SCI). Through the implementation of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, we gained valuable insights into the correlation between the microenvironment's heterogeneity and treatment efficacy. Concurrently, the molecular subtypes of genes with cell differentiation in SCI samples were screened in the dataset. Subsequently, we verified the correlation between the two molecular subtypes and immune microenvironment and the expression levels of corresponding immune checkpoints. In addition, we identified three immune markers, namely C1qa, Lgals3 and Cd63, through the screening of differentially expressed genes of normal cells and SCI cells, as well as the distinctively expressed genes in the two molecular subtypes. These markers were further validated through least Absolute contraction and Selection operator (LASSO) regression analysis, in conjunction with differentiation-related gene (DRG) crossover validation. Finally, we conducted verification experiments to assess the expression levels and diagnostic value of these three immune markers. In conclusion, our study successfully developed three corresponding biomarkers based on differentiation-related markers from single-cell RNA sequencing data of spinal cord injury immune cells in spinal cord injury. This research enhances our understanding of the cellular composition and underlying mechanisms of immune infiltration after spinal cord injury and presents a promising avenue for the implementation of precise immunotherapy strategies.

2. Materials and methods

2.1. Data acquisition, collation and processing

Derived from the GSE159638 dataset based on the GPL24247 platform in the GEO database (https://www.ncbi.nlm.nih.gov/geo/), we collected a total of 19,083 single-cell RNA sequencing (scRNA-seq data) from 2 laminectomy-only control samples and 14 SCI samples. Among the SCI samples, 7 were obtained at 3 days post-SCI, 5 were obtained at 10 days post-SCI, and 2 were obtained at 21 days post-SCI. Each sample consisted of tissue from 1 to 3 mice. In addition, we also incorporated transcriptomes data from the GSE18179 dataset, which utilized the GPL81 platform in the GEO database. This dataset provided transcriptomes from 32 SCI sample. The "PercentageFeatureSet" function was employed to assess the proportion of mitochondrial genes, and it was observed that both samples were excluded due to an increase number of mitochondrial genes. This observation implies a potential lack of cell viability. Subsequently, the single-cell RNA sequencing data underwent filtering and normalization utilizing the LogNormalize technique. The analysis focused on the top 1500 genes displaying high variability. It is noteworthy that all computational analyses conducted in this study were implemented using the R 4.2.0 programming language (https://www.r-project.org/).

2.2. Principal component analysis and cell annotation

To minimize redundancy, the authors applied principal component analysis (PCA) to identify the dimensions that exhibited significant separation. Then, the T-distributed stochastic neighbor embedding algorithm was employed to reduce the dimensions of the first 15 principal components and generate principal clusters. Marker genes for each cluster were identified using the "pheatmap v1.0.12″ package, with criteria such as log2 [fold-change (FC)] > 1 and false discovery rate (FDR) < 0.05. To annotate the clusters into specific cell types, the authors utilized the "SingleR v2.2.0″ package which matched the clusters with nine predefined cell types based on their respective marker genes. This approach aims to minimize redundancy in the analysis description.

2.3. Single-cell trajectory analysis and differentiation-related gene identification

Pseudotime and trajectory analyses of the cells were conducted using the “Monocle v2.5.4” package. Specifically, genes with $|\log2$ fold-change (FC)$| > 0.5$ and false discovery rate (FDR) < 0.05 were considered as intracellular DRGs,.

2.4. GO and KEGG enrichment analyses

The software packages "clusterProfiler v3.16.1″, "enrichplot v1.10.2″, "org.Hs.eg.db v3.12.0″, and "ggplot2 v3.4.0″ were employed for performing GO function analysis and KEGG pathway enrichment analysis of the DRGs in the three subsets. Statistical significance was determined at a threshold of $P < 0.05$ to exclusion of redundant findings and enhance the rigor of the analysis.

2.5. DRG-based molecular subtypes

The "ConsensusClusterPlus v1.54.0″ package was employed to conduct consensus clustering and determine the number of unsupervised subtypes. To identify stable subtypes, the K-means algorithm and the cumulative distribution function (CDF) were applied with 50 iterations and a maximum of 9 subtypes. Through this iterative process, the optimal number of subtypes were determined through this process.

2.6. Analysis of immune microenvironment, infiltrating immune cells, and immune checkpoints

The ESTIMATE algorithm v1.0.13 was utilized to calculate the immune-to-stromal ratio within the immune microenvironment for each molecular subtype. Additionally, the CIBERSORT algorithm was applied to estimate the abundance of 22 infiltrating immune cells in each sample. Furthermore, differential expression analysis was conducted to identify and evaluate the expression patterns of 39 validated immune checkpoints. By utilizing this comprehensive approach, redundancy in the analysis was minimized.

2.7. DEGs identification and venn diagramming

The limma v3.46.0 package in R was utilized to assess the differential expression of genes (DEGs) in the ScRNA. Specifically, DEGs were identified between normal cells and SCI cells from the GSE159638 dataset. Furthermore, subgroup DEGs were identified between subgroups from GSE159638 dataset. The threshold conditions for differential gene expression screening were set as "adjusted $P < 0.05$ and log2 (FC) > 1 or log2 (FC) < -1". To identify genes specifically associated with SCI, a Venn diagram analysis was performed (http://vip.sangerbox.com/home.html).

2.8. Identification of biomarkers in SCI

The LASSO algorithm from the glmnet v4.1 package in R software was employed to select significant genes from the identified DEGs. Subsequently, the expression levels of these hub genes were assessed using box plots in both SCI and control samples. To determine the diagnostic value of the identified genes, receiver operating characteristic (ROC) curves were plotted. These curves provided valuable insights into the sensitivity and specificity of the genes in distinguishing between SCI and control samples. The genes selected through this process hold significant potential for improving the diagnosis of SCI.

2.9. Mice spinal cord injury model

In this study, female C57BL/6 mice aged between eight and ten weeks were utilized. The mice were purchased from Vital River Laboratory Animal Technology Co., Ltd. The surgical procedures were performed using the contusion model, and the mice were anesthetized with isoflurane (RWD, R510-22, Guangdong, China). After making a longitudinal incision through the skin and fascia, the muscles were carefully separated, exposing the T7-T9 spinous processes. A T8 laminectomy was performed to fully expose the spinal cord. Fixators were utilized to stabilize the mouse T7 and T9 tissues by attaching them to the spinal processes. Subsequently, a mild spinal cord contusion model was induced using the NYU Impactor-III (WM Keck, USA) with a force of 5 g and a displacement of 12.5 mm. Once the successful contusion was confirmed, the muscles, fascia, and skin were sutured. Post-spinal cord injury, manual bladder emptying was performed twice daily. Urinary tract infections were treated with intramuscular injections of cefuroxime sodium (6 mg/mL, dissolved in saline; Hongtu, Nanjing, China).

2.10. qRT-PCR

Spinal cord tissues from 6 mice were collected at 3 days post-surgery in both the sham operation group and the spinal cord injury group. Total RNA was extracted from the spinal cord tissues and cDNA synthesis was performed following the producer's protocol (Takara Biotech Co, Ltd, Beijing, China). Quantitative real-time PCR (qRT-PCR) was conducted on an Applied Biosystems 7500 instrument (Thermo Fisher Scientific, China, Shanghai) utilizing Real-Time PCR Mix (Cowin Biotech, Taizhou, Jiangsu Province, China). The $2-\Delta\Delta Ct$ method was employed to determine the relative gene expression levels normalized to GAPDH and differentially expressed genes (DEGs). The experiment was independently repeated three times. The RNA sequence fragments are presented in Table 1.

2.11. Statistical analysis

The differences in the expression status of the 59 m7G, m6A, m5C, and m1A regulators between the healthy and SS samples were compared using the Wilcoxon test. The $\chi 2$ test and T-test/variance analysis were separately used to compare the distribution of dichotomous variables and continuous variables. All statistical analyses were performed using R, Perl, and Graphpad software. A significance level of $P < 0.05$ was considered statistically significant.

3. Results

3.1. Quality control of scRNA-seq data and principal component analysis

A total of 19083 cells from the GSE159638 dataset were included in the analysis after quality control, filtering, and batch effect correction of scRNA sequence data (Fig. 1A). The number of detected genes showed a positive correlation with sequencing depth ($R = 0.94$, Fig. 1B). After filtering out low-quality cells and normal cells, gene expression data from 17196 SCI cells were further analyzed. Among all the samples, 22,133 genes demonstrated variations, and the top 1500 genes with the most notable variations were selected for additional analysis (as depicted in Fig. 1C). To reduce the dimensionality of the scRNA-seq data, principal component analysis (PCA) was performed, as demonstrated in Fig. 1D. The top 10 genes, including Cxcl10, Ccl5, S100a9, Ccr7, and Hist1h2ap were identified as having significant differences among the cell samples. A Jack Straw Plot was then employed to visualize the distribution of p-values for each principal component (PC), as shown in Fig. 1E. Based on a significance threshold of $P < 0.0001$, the top 15 PCs with significant differences were selected further analysis.

3.2. Three subsets of SCI were identified by single cell trajectory analysis

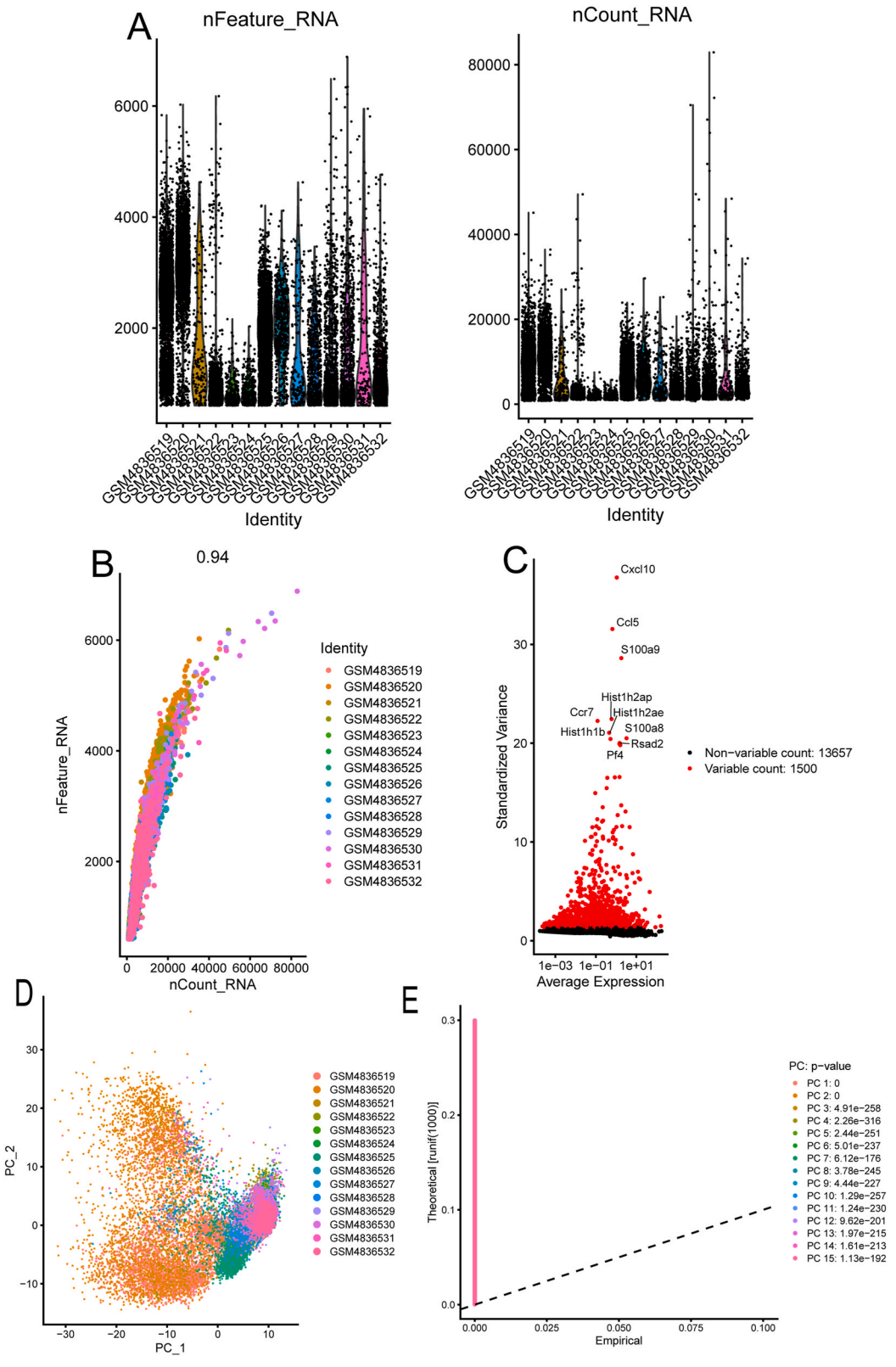
The tSNE algorithm was applied to aggregate the 17196 SCI data into 12 clusters (Fig. 2A). Through the analysis, a total of 2569 marker genes were identified and visualized in a heatmap. The SCI data from 17196 observations were analyzed using the tSNE algorithm, resulting in the identification of 12 distinct clusters, as depicted in Fig. 2A. Furthermore, to gain insights into the gene expression patterns with each cluster, a heatmap was generated to display the 2569 marker genes associated with each cluster, as shown in Fig. 2B. By annotating the 12 clusters, it was possible to assign them into 6 cell types. Notably, these findings highlighted the presence of a diverse and abundant immune microenvironment within the SCI samples, as depicted in Fig. 2C. Following trajectory and pseudotime analysis, the results showed that the three branches of differentiated cells exhibited distinct variation along the pseudotime axis (Fig. 2D). Pseudotime analysis, employing an unsupervised machine learning algorithm, enables the inference and simulation of cell differentiation progression and order using single-cell trajectories, types, and clusters, allows for the detection of changes in the immune microenvironment based on corresponding cell types and time series (as illustrated in Fig. 2E and F). In summary, the SCI immune microenvironment comprises various types of immune cells as the primary and foundational components. This is followed by an increase in the number of stem cell-like cells, which may act as precursors for spinal cord cells. Over time, there is a gradual increase in the proportion of spinal cord cells, indicating the developmental process of the spinal cord. Through this analysis, a set of 404 differentially expressed genes (DEGs) were identified for each trajectory, which were classified as differentiation-related genes (DRGs), effectively reducing redundancy in the dataset.

3.3. GO and KEGG enrichment analyses of DRGs in three subsets

To investigate the roles of DRGs in the three subsets, the researchers conducted KEGG pathway enrichment and GO analyses. The GO analysis revealed that the DRGs were primarily involved in regulating immune system processes across all three subsets. Specifically, in subset I, the DRGs were involved in the regulation of myeloid leukocyte migration and mononuclear cell migration. In subset

Table 1
The sequence fragments of RNA.

	Forward	Reverse
C1qa	5′-AAAGGCAATCCAGGCAATATCA-3′	5′-TGGTTCTGGTATGGACTCTCC-3′
Cd63	5′-GAAGCAGGCCATTACCCATGA-3′	5′-TGACTTCACCTGGTCTCTAAACA-3′
Lgals3	5′-TGCTGGTTCCAGGGACTCAA-3′	5′-CCACCGGCCTCTGTAGAAGA-3′



(caption on next page)

Fig. 1. Quality control and screening of scRNA-seq data. (A) The number of genes identified in 14 SCI samples and sequencing depth in each sample. (B) Relevance analysis between total intracellular sequences and sequencing depth(R = 0.94). (C) Volcano plot demonstrated genes fluctuating in all samples. The top 1500 genes marked in red had remarkable variations, and the names of the top 10 genes are displayed. (D,E) PCA was conducted to identify the significantly available dimensions of data sets with estimated P value.Classification of cells by PCA.

II, the DRGs were related to the positive regulation of T cell activation, lymphocyte activation, and cytokine production. In subset III, the DRGs were primarily enriched in the negative regulation of leukocyte differentiation, hemopoiesis, and endopeptidase activity (Fig. 3A–C). The KEGG pathway enrichment analysis revealed that the three subsets were associated with infection, antigen processing and presentation, hematopoietic cell lineage, as well as various immune-related and autoimmune diseases (Fig. 3D–F).

3.4. The DRG-based molecular subtypes were identified from SCI samples in the GSE18179 dataset

Using the identified 404 marker genes as differentiation-related genes (DRGs), the researchers obtained transcriptomes of 32 SCI samples from the GSE18179 dataset. They performed DRG-based consensus clustering analysis. With a clustering threshold of Kmax equal to 9, two molecular subtypes were identified in SCI samples based on the DRG intersection of three branches (Fig. 4A–C). Additionally, the expression patterns of subsets I/II/III of DRG were found to be similar in subtype I/II (C1/2), indicating that subsets I/II/III were respectively included in subtype I/II (C1/2). Therefore, it can be concluded that subtype I/II (C1/2) was composed of subsets I/II/III (Fig. 4D–F).

3.5. Analysis of immune microenvironment, infiltrating immune cells, and immune checkpoints in two molecular subtypes

In this study, the researchers utilized the ESTIMATE algorithm to evaluate the immune/stromal components of each molecular subtype. Prior to that, they applied the R homologene function to map the 8642 mouse (Mus musculus) genes in the GSE18179 dataset to their corresponding human homologs with maximum precision. Duplicate mappings were removed, resulting in an expression matrix consisting of 7743 human homologs for subsequent analysis. In this study, the researchers evaluated the immune and stromal components of the immune microenvironment using the ImmuneScore/StromalScore respectively. These scores represent the abundance of immune and stromal cells in each sample. Additionally, the ESTIMATEScore, which is the sum of the ImmuneScore and StromalScore, was calculated as a composite measure representing the overall immune microenvironment. The ImmuneScore, StromalScore, and ESTIMATEScore demonstrated significantly higher values (all $p < 0.001$) in cluster C2 compared to cluster C1, indicating a greater abundance of immune and stromal cells (Fig. 5A–C). Conversely, cluster C1 exhibited a lower proportion of immune cells and stromal cells. To further investigate and confirm the relationship between molecular subtypes and immune microenvironment, we employed the CIBERSORT algorithm to estimate the abundance of 22 different infiltrating immune cells in each sample. The resulting data were graphically represented for each of the two molecular subtypes (Fig. 5D). Our analysis revealed notable distinctions in immune cell infiltration between the two subtypes. The C2 cluster exhibited a higher abundance of resting CD4 memory T cells, plasma cells, and activated mast cells, while the C1 cluster demonstrated a higher enrichment of follicular helper T cells, regulatory T (Tregs) cells, resting NK cells, monocytes, M0 macrophages, resting dendritic cells, and activated mast cells, as shown in Fig. 5E. Additionally, we observed differential expression levels of immune checkpoints. TNFSF4, LAG3, JAK1, IL12B, IL12A, IFNG, ICOS, and CD40LG exhibited higher expression in the C2 cluster, whereas PTPRC, LDHA, JAK2, and B2M displayed higher expression in the C1 cluster (Fig. 5F).

3.6. Identification DEGs in GSE159638 dataset and GSEA enrichment analyses of DEGs in GSE18179 dataset

We identified a total of 392 DEGs between 1887 normal cells and 17,196 SCI cells from the GSE159638 dataset, which were classified as scRNA DEGs (Fig. 6A). Additionally, we identified 100 DEGs between cluster C1 and cluster C2 from the GSE18179 dataset, which were considered as subgroup DEGs (Fig. 6B). GSEA results revealed significant associations between the C2 phenotype and several biological processes, including focal adhesion, natural killer cell mediated cytotoxicity, ECM receptor interaction, primary immunodeficiency, pentose and glucuronate interconversions, and cytokine-cytokine receptor interaction (Fig. 6C).

3.7. Acquisition of key biomarkers for SCI

By intersecting the differentiation-related genes (DRGs) with scRNA DEGs and subgroup DEGs, a total of 11 related genes identified: C1qa, Iqgap1, Lgals3, pro1, Ahnak, Gsn, anx1, Cd63, Emp1, S100a8 and Lpl(Fig. 7A). Subsequently, these 11 genes were analyzed by LASSO regression to identify gene markers with high diagnostic value using a machine learning approach. Following LASSO coefficient profiling (Fig. 7B) and validation (Fig. 7C), three candidate markers, namely C1qa, Lgals3 and Cd63, were identified. The expression patterns of the three essential biomarkers demonstrated in Fig. 7D–F revealed variations among different cell subsets. Lgals3 exhibited widespread expression across most cells, while C1qa showed high expression levels in neutrophils and dendritic cells. Coversely, Cd63 displayed high expression in neutrophils, dendritic cells, and platelets, distinguishing it from other genes. These findings suggest that these biomarkers may serve distinct roles in different cell types.

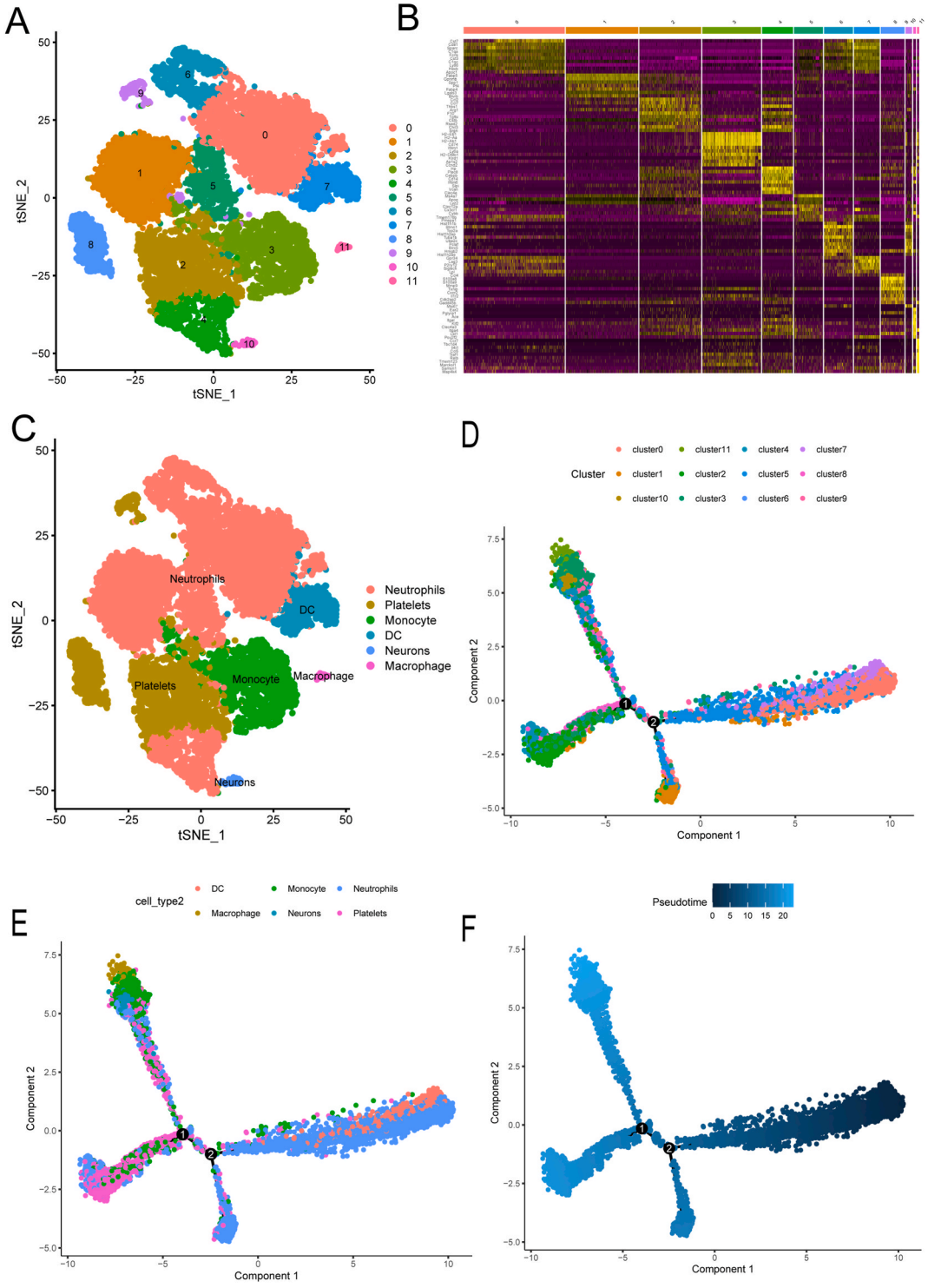


Fig. 2. Identification of SCI subsets by the single-cell trajectory analysis. (A) tSNE was applied for clustering of SCI cells. (B) Heatmap of marker genes in each cluster.Up-regulated genes were marked with yellow and down-regulated genes were marked with purple. (C) Cell type labeling of each cluster.On the basis of the marker genes, 12 clusters were annotated into 6 cell types. (D) Single-cell pseudotime analysis of three subsets. Dark blue indicated an earlier time. (E, F) Trajectory analysis of cell types and clusters. The different colors' dots denoted the corresponding clusters or cell types, which were arrayed on the pseudotime branches.

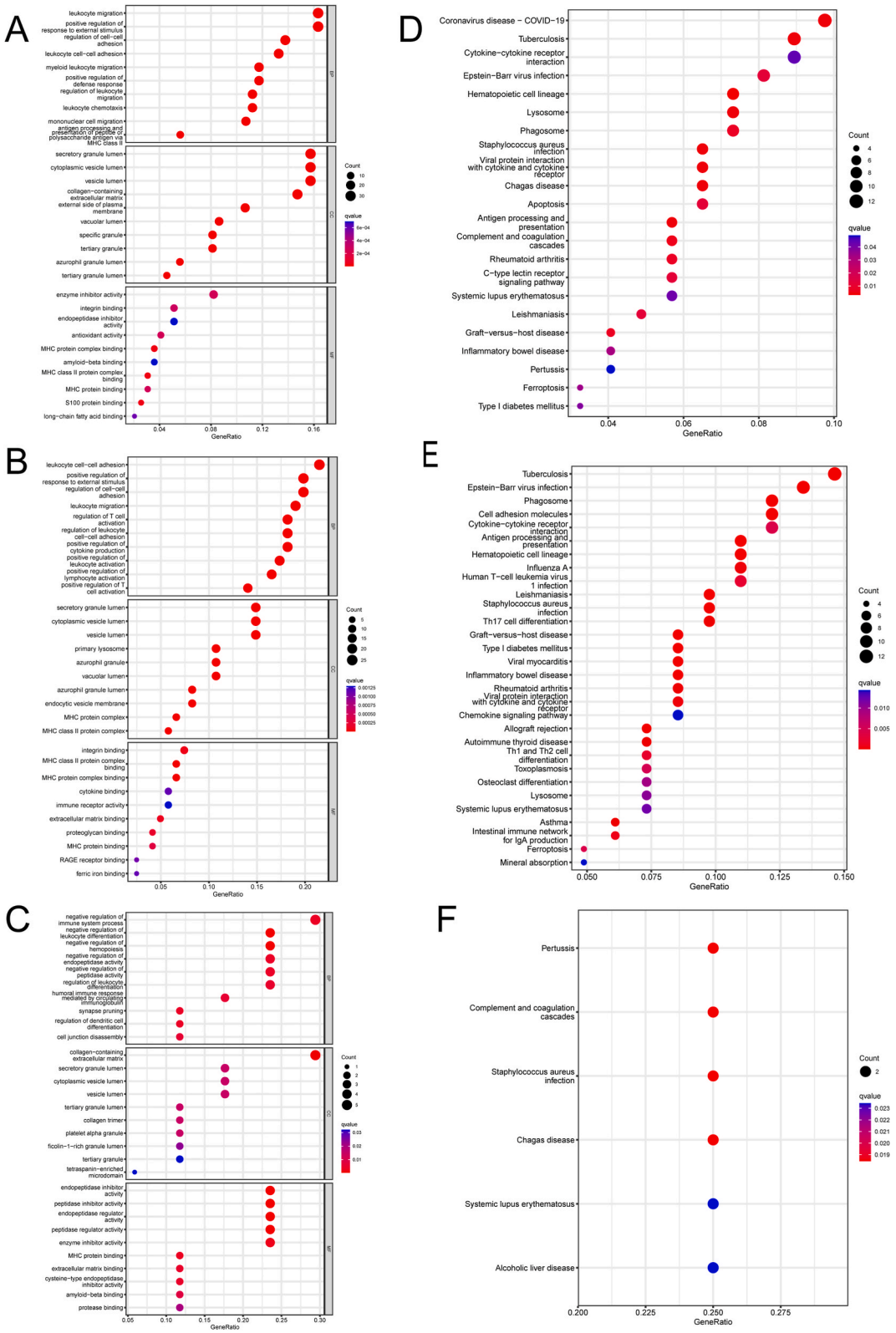


Fig. 3. GO and KEGG enrichment analyses of DRGs in three subsets. (A–C) GO functional analysis of DRGs in subsets I-III. (D–F) KEGG pathway enrichment analysis of DRGs in subsets I-III.

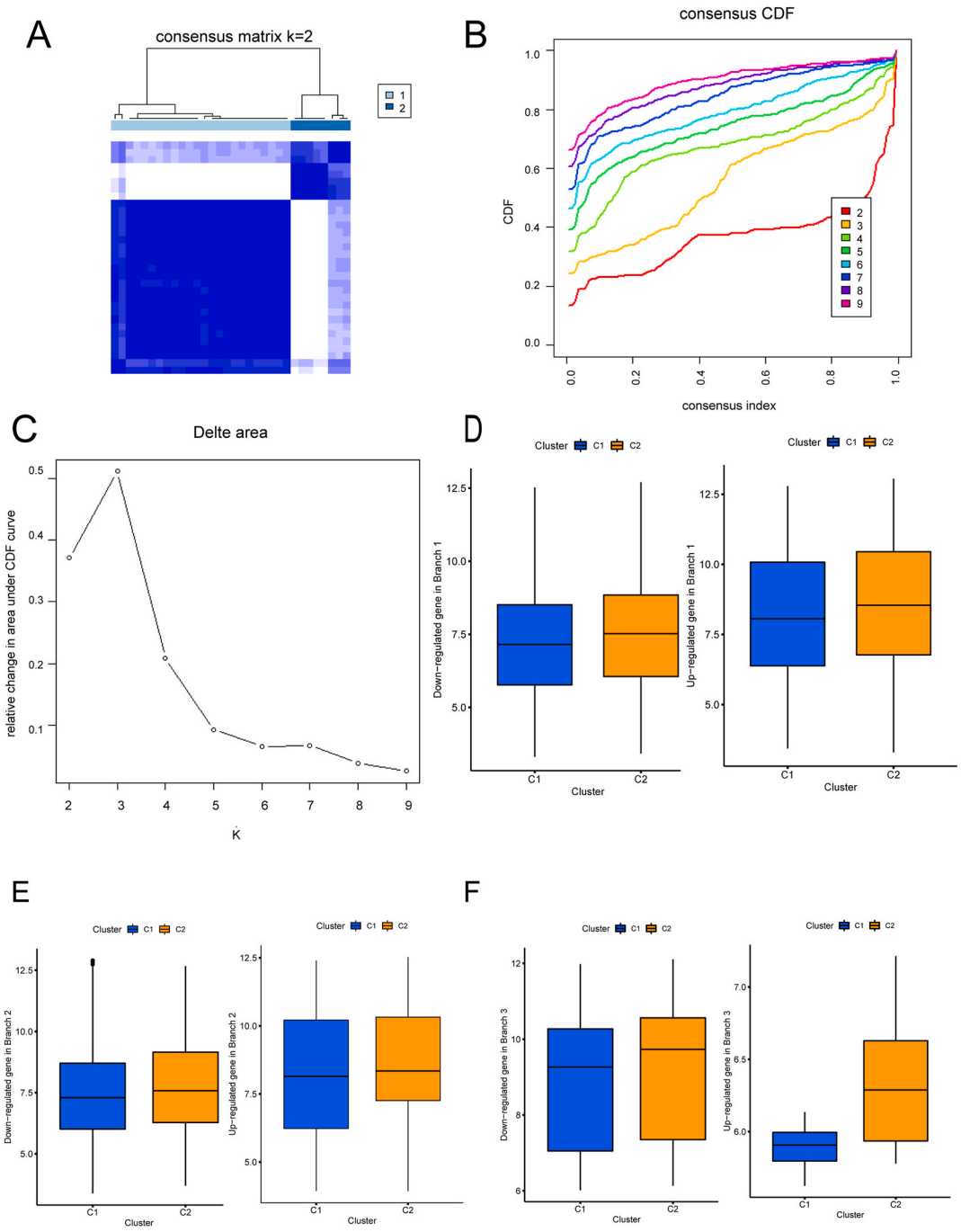


Fig. 4. DRG-based molecular subtypes were identified from SCI samples. (A–C) The outcomes of consensus clustering analysis for SCI samples on the basis of DRGs. (D-F)The up/down-regulated DRGs in different state 1 to 3 displayed similar expression patterns in C1 and C2 clusters, respectively.

3.8. Identification of key biomarkers expression levels and diagnostic value

To validate the expression levels of the three hub genes, box plots were employed. As illustrated in Fig. 6A, the expression levels of C1qa, Cd63, and Lgals3 were found to be significantly higher in SCI samples as compared to controls ($P < 0.001$ for C1qa and Lgals3, and $P < 0.05$ for Cd63) (Fig. 8A–C). To evaluate the diagnostic potential of these three hub genes for SCI, ROC curve analysis was conducted, and the area under the curve (AUC) values were determined. Two of the hub genes exhibited AUC values exceeding 0.77, while Cd63 showed an AUC value of 0.651 (Fig. 8D–F).

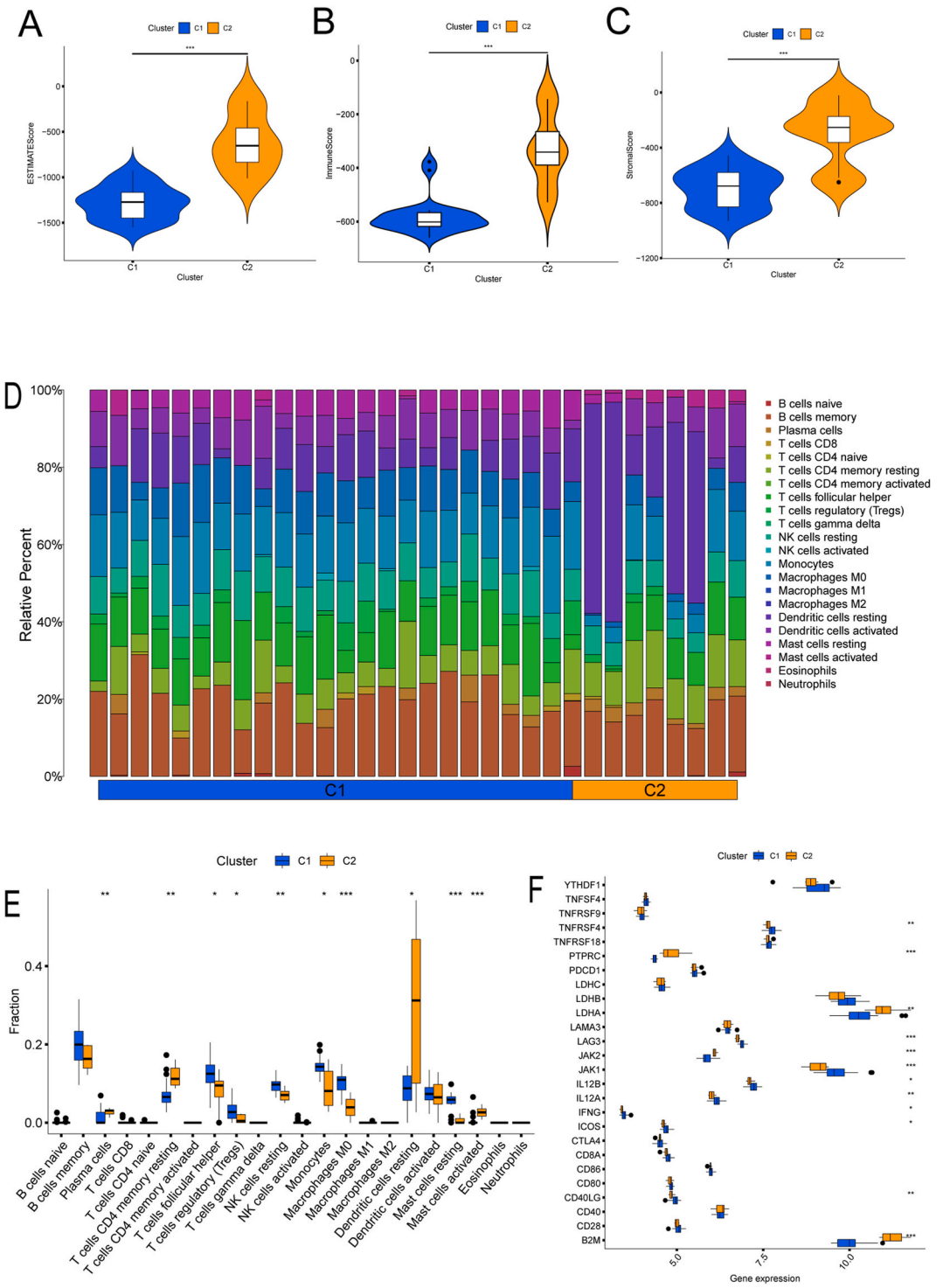


Fig. 5. Analysis of immune microenvironment and infiltrating immune cells in two molecular subtypes. (A–C) ImmuneScore, StromalScore and ESTIMATEScore in two molecular subtypes. (D) The contents of 22 infiltrating immune cells in each sample are displayed according to two molecular subtypes. (E) Differential expression analysis of infiltrating immune cells in two molecular subtypes. (F) Comparisons of the expression levels of immune checkpoints. (*P < 0.05, **P < 0.01, ***P < 0.001).

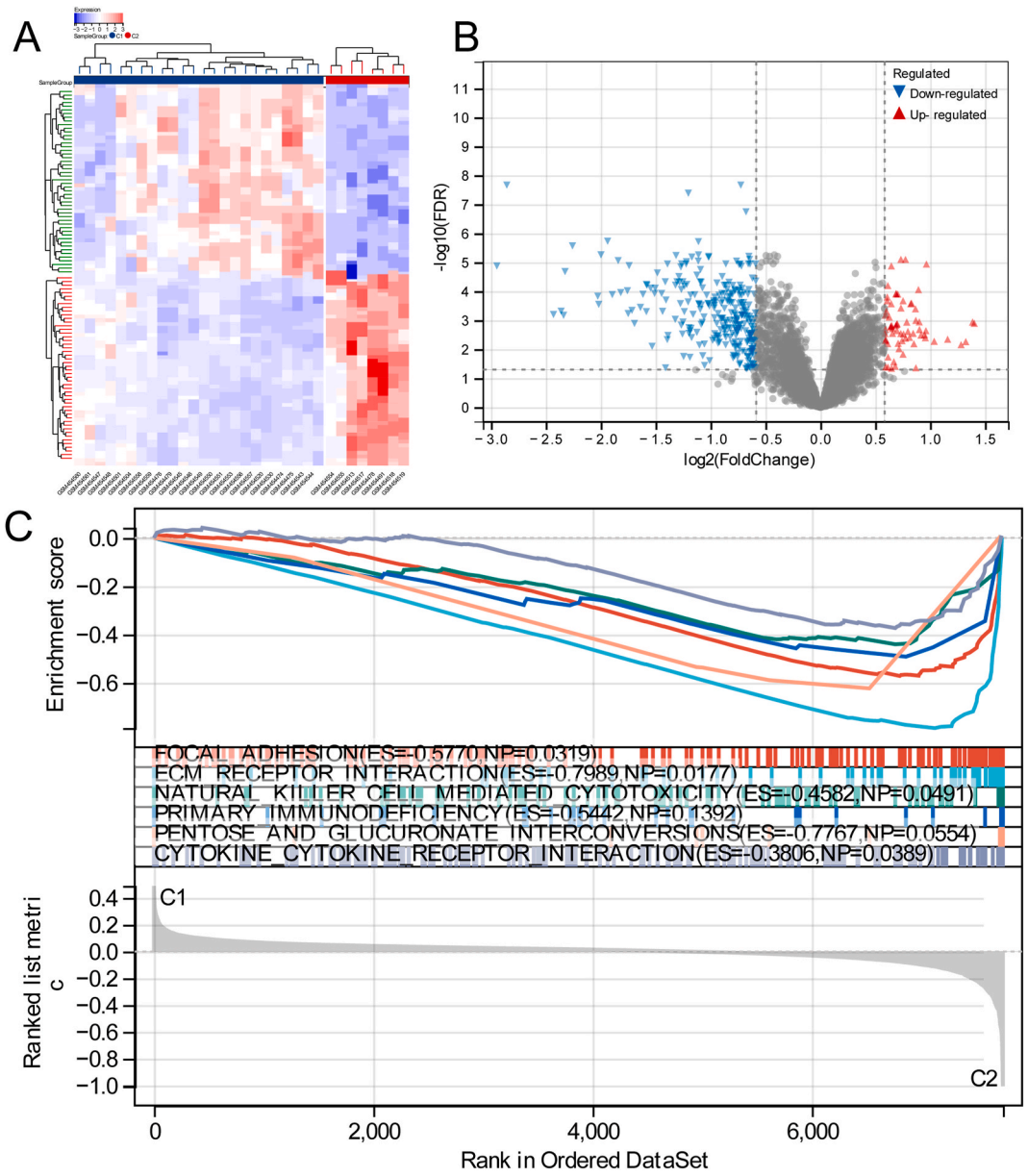


Fig. 6. Differential mapping and functional enrichment analysis of two molecular subtypes in SCI. (A) Thermal plots of scRNA DEGs. (B) Volcano plot of subgroup DEGs. (C) The remarkable enriched signalling pathways between two subgroups in the GSEA of KEGG.

3.9. Expression level of hub genes

To further validate the accuracy of our findings, we identified the expression of 3 hub genes by qRT-PCR. The PCR results confirmed that the expression of these 3 hub genes, which were significantly related to SCI, was significantly up-regulated in the SCI group. These findings were consistent with our previous analysis(Fig. 9A–C).

4. Discussion

The severe, destructive, and persistent inflammatory response after spinal cord injury can exacerbate local tissue damage and lead to the development of syringomyelia, characterized by the infiltration of phagocytic macrophages at the injury site. In recent years, there has been a growing emphasis on understanding the immune microenvironment after spinal cord injury, leading to the exploration of various immunotherapeutic approaches for spinal cord injury repair, demonstrating promising therapeutic prospects [24–26]. The efficacy of therapeutic interventions for spinal cord injury (SCI) often fall below expectations, primarily due to factors

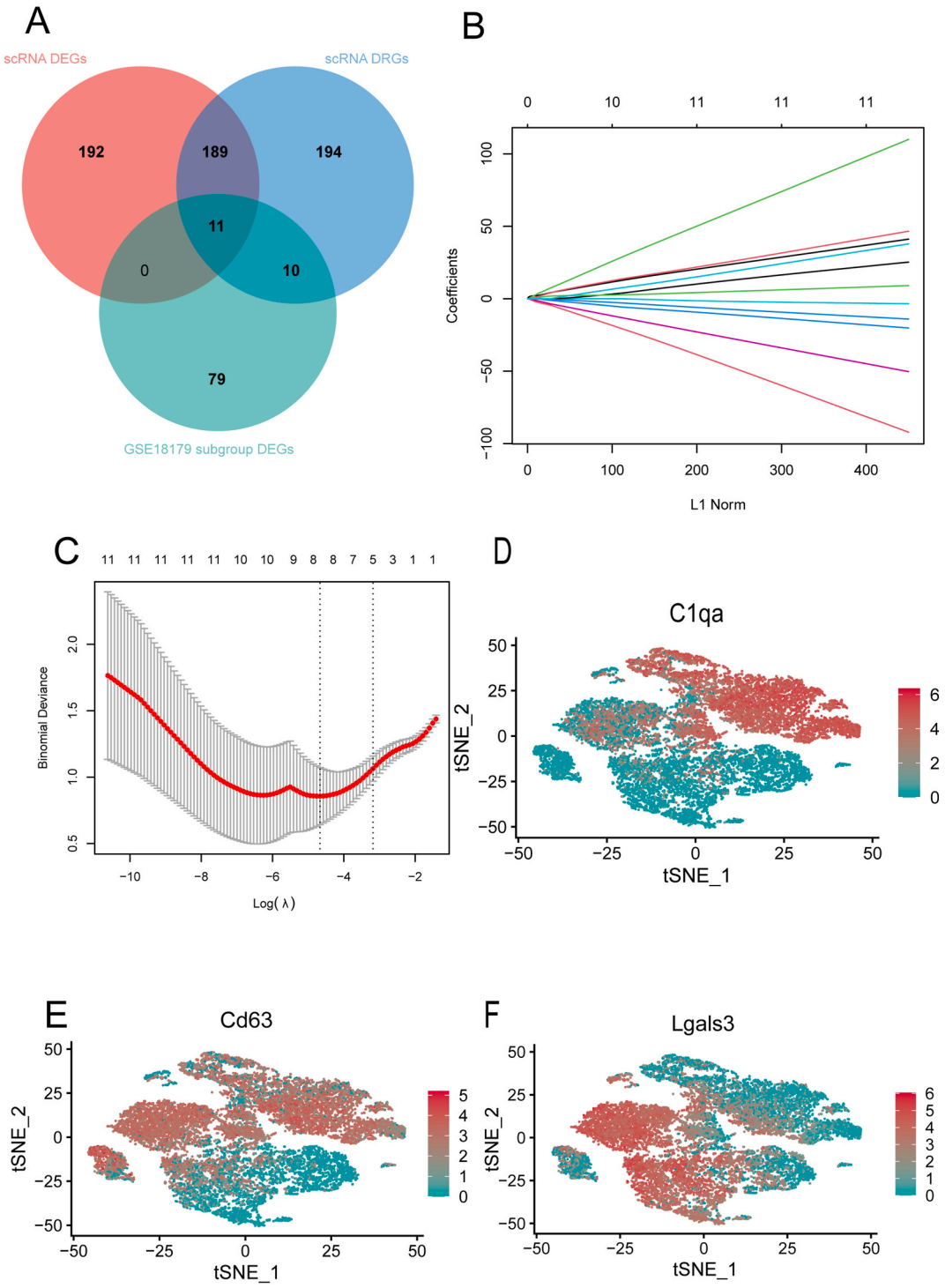


Fig. 7. Screening of Hub Genes. (A) Venn diagram for intersections among DRGs, scRNA DEGs and subgroup DEGs. (B) Partial likelihood deviance with changing of log (λ) plotted by LASSO regression in 10-fold cross-validations. Dotted vertical lines were drawn at the optimal values using the minimum criteria (lambda.min) and 1 standard error of the minimum criteria (1-SE criteria). (C) The LASSO coefficient profiles for five hub genes in the 10-fold cross-validation. (D–F) The expression patterns of the 3 hub genes in the subgroups of the single-cell RNA-seq data.

such as the diverse immune microenvironment following SCI, among others. In addition, evidence suggests that immune checkpoints, which serve as major targets for immunotherapy, may play a role in the regulation the activity of microglia and infiltrating immune cells [27]. Consequently, the regulation of immune checkpoints and the development of corresponding immune markers are conducive

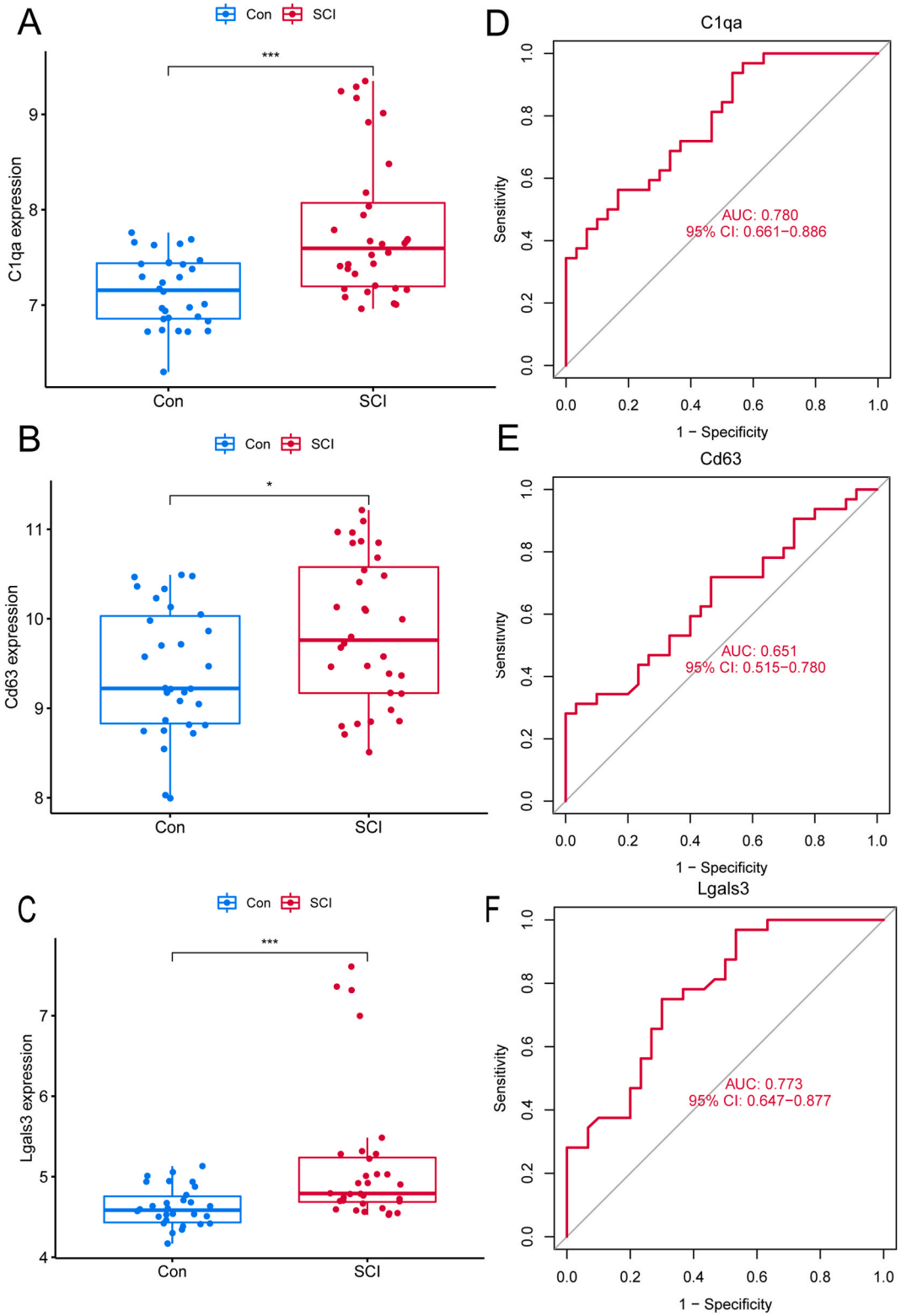


Fig. 8. Identification of hub genes in the gene expression level and the diagnostic value in SCI. (A–C) Expression box plots of candidate genes. (D–F) ROC curve evaluation of candidate genes. (*P < 0.05, ***P < 0.001).

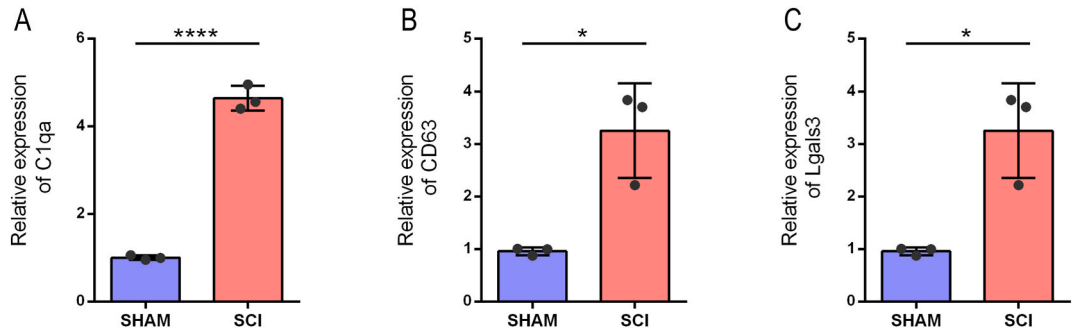


Fig. 9. Expression level of hub genes in sham control and SCI mice.(A) Expression level of C1qa.(B) Expression level of CD63.(C) Expression level of Lgals3. (*P < 0.05, ****P < 0.0001).

to understand the heterogeneity of immune infiltration, facilitating the diagnosis of SCI and accurate immunotherapy.

ScRNA-seq technology has proven useful in investigating the specific cell types present and differentiation states following SCI. This innovative approach enables the identification and characterization of specific cell types and their transcriptional profiles within the local immune microenvironment. However, the current prognosis for patients with acute spinal cord injury primarily relies on the assessment of residual neurological function based on lesion characteristics [28,29]. Existing immune markers for the prognosis of spinal cord injury mainly stem from peripheral blood tests [30–32]. However, there remains a paucity of research exploring potential immunotherapy targets and biomarkers for SCI prognosis using local tissue scRNA sequencing data. Such investigations hold the promise of providing a novel exploration perspective for SCI treatment, which also emphasizes the need to identify prognostic biomarkers specific to SCI.

In this study, we leveraged the data from the SCI scRNA-seq, through cell differentiation path to explore the impact of the spinal cord injury immune microenvironment by exploring a broader range of genes associated with cell differentiation. The genetic heterogeneity of the differentiated spinal cord was assessed by employing the ESTIMATE algorithm to calculate the sample matrix, immune. The West Burr sorting algorithm was utilized to conduct a further analysis to verify the correlation between molecular subtypes and the immune microenvironment. Finally, we successfully identified immune checkpoints that exhibited higher expression in the C1 subtype (PTPRC, LDHA, JAK2, B2M). the C2 subtype displayed elevated expression levels of immune checkpoints (TNFSF4, LAG3, JAK1, IL12B, IL12A, IFNG, ICOS, CD40LG). Then, the key SCI biomarkers (C1qa, Lgals3 and Cd63) were obtained by employing the LASSO regression. The expression levels of these biomarkers were then verified through boxplot and ROC curve analysis, allowing for the evaluation of their diagnostic specificity and sensitivity. Finally, qRT-PCR was applied to compare the differences in expression levels of three hub genes between SCI group and the sham group.

Our study has identified C1qa, Lgals3, and Cd63 as important biomarkers for SCI based on our findings, with their expression levels showing a significant increase after SCI compared to the sham group. In terms of their expression patterns, C1qa exhibited high expression in neutrophils and dendritic cells, Lgals3 displayed widespread expression across most cells, and Cd63 was highly expressed in dendritic cells, neutrophils and platelets. Notably, C1qa overproduction and activation trigger the classical complement pathway, which mediates microglial phagocytosis [2][33]. Geoffrey T. Norris et al. conducted a study that demonstrated the involvement of complement C1q in microglial phagocytosis of neuronal and myelin debris [3][34]. Research has also indicated that the complement system mediates the loss of neuronal synapses in multiple sclerosis (MS), with the difference that C1q is primarily derived from neurons in MS [4][35]. In a recent study conducted by Andrew N Stewart et al., it was found that showed that only the complement protein C1qa showed a significant increase in males among the genes differentially expressed between genders in MDM after SCI. In addition, the expression levels of C1qa, cPLA2 and CD86 genes were found to be expressed at higher levels in microglia compared to those in MDM[5] [36]. These findings suggest that the upregulation of C1qa markers in microglia is a common occurrence after spinal cord injury and are more applicable to the male population. Similarly, Kimberly R Byrnes et al. confirmed that the expression of Lgals3 (galectin-3) was significantly up-regulated at both 28 and 6 months after SCI compared to the sham group, and immunolabeling confirmed that galectin-3 was predominantly expressed in Iba1+ cells, which was inconsistent with our analysis results[6] [37]. Additionally, Zhouliang Ren et al. provided data that supports the upregulation of Gal-3 expression levels after SCI and its crucial role in regulating the severity of SCI neuroinflammation through enhancing the activation of the ROS/TXNIP/NLRP3 signaling pathway[7] [38]. Therefore, Gal-3 is a promising biomarker for spinal cord injury, and its deficiency attenuates neuroinflammation via ROS/TXNIP/NLRP3 signaling

2 [33].The number of the reference needs to be updated.
3 [34].The number of the reference needs to be updated.
4 [35].The number of the reference needs to be updated.
5 [36].The number of the reference needs to be updated.
6 [37].The number of the reference needs to be updated.
7 [38].The number of the reference needs to be updated.

pathway. CD63, currently recognized as an exosome marker, plays a significant role in facilitating exosomes secretion from cells [8][39]. In addition, recent studies have also confirmed the upregulation of CD63 expression in both plasma and spinal cord tissues after SCI[9] [40]. The findings of Sara Salvany et al. suggest the presence of abundant extracellular vesicles within the perineuronal space after axotomy, accompanied by an increased expression of phospho-mixed lineage kinase domain-like protein. The study reports an increased expression of extracellular vesicle markers, including CD9, CD63 and flotillin at later stages. Activated microglia phagocytosed cell fragments, and synaptic boutons showed positive C1q immunoreactivity, indicative of complement-mediated conditioning that may also contribute to microglia-mediated synaptic disruption [10][41]. These findings suggests a role for CD63 as an early inflammatory marker in spinal cord injury, implicating neuronal fragmentation and the release of extracellular vesicles.

The bioinformatics analysis employed in identifying SCI biomarkers revealed higher AUC values for the three genes, indicating increased sensitivity and accuracy in identifying SCI. These trends were further validated in PCR results. In addition, our study successfully screened out highly expressed immune checkpoints for the two molecular subtypes. This finding enhances our understanding of the heterogeneity of local immune infiltration after SCI and facilitates the exploration of potential potential therapeutic interventions by identifying candidate drugs. These results serve as a valuable reference for future immunotherapy approaches in the field of SCI.

The present study has several limitations that should be acknowledged. Firstly, the majority of our findings are based on bioinformatics analysis, and the samples used in this study were obtained from rodents rather than humans subjects. As such, further exploration and validation through clinical and relevant biological experiments are required. Secondly, the expression pattern of Lgals3, a key biomarker of SCI in our study, is known to be widely expressed in most cells. However, a study conducted by Geoffrey T. Norris and colleagues suggested that Lgals3 is mainly expressed in Iba1+ cells. Therefore, additional experiments are needed to verify the inconsistent mechanisms associated with Lgals3 expression. This discrepancy represents another limitation of our study.

5. Conclusions

In conclusion, this study conducted a comprehensive comparative analysis of single-cell sequencing data obtained from SCI samples and control samples in the GEO database.Through this analysis, our study revealed a diverse immune microenvironment in SCI samples, highlighting the potential significance of C1qa, Lgals3 and Cd63 as immune biomarkers for diagnosing SCI. Moreover, the identification of immune checkpoints corresponding to the two molecular subtypes suggests their potential as targets for immunotherapy to enhance SCI repair in future interventions.

Contributions

(I) Conception and design: Jun Shang; (II) Administrative support: Shiqing Feng (III); Provision of study materials:Chao Ma and Han Ding (IV); Collection and assembly of data: Guangjin Gu and Jianping Zhang; (V) Data analysis and interpretation: Han Ding and Zhijian Wei; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Ethical statement

The study was approved by the Experimental animal Welfare Ethics Committee of Tianjin Medical University General Hospita (No. IRB2022-DW-47). The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Author contribution statement

Chao Ma: Han Ding: Performed the experiments; Wrote the paper; Guangjin Gu: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper; Jianping Zhang: Min Wang: Ke Fang: Contributed reagents, materials, analysis tools or data; Wrote the paper; Zhijian Wei: Analyzed and interpreted the data; Wrote the paper; Shiqing Feng: Jun Shang: Conceived and designed the experiments; Wrote the paper.

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

[8] [39].The number of the reference needs to be updated.
[9] [40].The number of the reference needs to be updated.
[10] The number of the reference needs to be updated.

Acknowledgments

Funding:This work was supported by grants from the National Key R&D Program of China (2019YFA0112100) and the National Natural Science of China (81930070).

Abbreviations

SCI	spinal cord injury
PCA	principal component analysis
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
ROC	receiver operating characteristic curves
CNS	central nervous system
scRNA	single-cell RNA
LASSO	least Absolute contraction and Selection operator
DRG	differentiation-related gene
tSNE	T-distributed stochastic Neighbor Embedding
PCs	principal components
FDR	false discovery rate
CDF	cumulative distribution function
DEGs	differentially expressed genes
AUC	area under the curve
LDH	lactate dehydrogenase
JAK1	Janus kinase-1
ICOS	Inducible T-cell costimulatory
MS	multiple sclerosis

References

[1] B. Fan, et al., Microenvironment imbalance of spinal cord injury, Cell Transplant. 27 (6) (2018) 853–866.
[2] B. Fan, Z. Wei, S. Feng, Progression in translational research on spinal cord injury based on microenvironment imbalance, Bone Res 10 (1) (2022) 35.
[3] H. Huang, et al., Clinical neurorestorative therapeutic guidelines for spinal cord injury (IANR/CANR version 2019), J Orthop Translat 20 (2020) 14–24.
[4] S.L. James, et al., Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016, Lancet Neurol. 18 (1) (2019) 56–87.
[5] M.D. Stillman, et al., Complications of spinal cord injury over the first year after discharge from inpatient rehabilitation, Arch. Phys. Med. Rehabil. 98 (9) (2017) 1800–1805.
[6] C.S. Ahuja, et al., Traumatic spinal cord injury, Nat. Rev. Dis. Prim. 3 (2017), 17018.
[7] D.C. Shields, A. Haque, N.L. Banik, Neuroinflammatory responses of microglia in central nervous system trauma, J. Cerebr. Blood Flow Metabol. 40 (1_suppl) (2020) S25–S33.
[8] A.D. Greenhalgh, S. David, Differences in the phagocytic response of microglia and peripheral macrophages after spinal cord injury and its effects on cell death, J. Neurosci. 34 (18) (2014) 6316–6322.
[9] D.J. Hellenbrand, et al., Inflammation after spinal cord injury: a review of the critical timeline of signaling cues and cellular infiltration, J. Neuroinflammation 18 (1) (2021) 284.
[10] X. Freyermuth-Trujillo, et al., Inflammation: a target for treatment in spinal cord injury, Cells 11 (17) (2022).
[11] M.B. Bracken, Steroids for acute spinal cord injury, Cochrane Database Syst. Rev. 1 (2012) CD001046.
[12] C.A. Bowers, B. Kundu, G.W. Hawryluk, Methylprednisolone for acute spinal cord injury: an increasingly philosophical debate, Neural Regen Res 11 (6) (2016) 882–885.
[13] M. Stenudd, H. Sabelstrom, J. Frisen, Role of endogenous neural stem cells in spinal cord injury and repair, JAMA Neurol. 72 (2) (2015) 235–237.
[14] L. Fan, et al., Exosomes-loaded electroconductive hydrogel synergistically promotes tissue repair after spinal cord injury via immunoregulation and enhancement of myelinated axon growth, Adv. Sci. 9 (13) (2022), e2105586.
[15] C. Zhao, et al., Chronic spinal cord injury repair by NT3-chitosan only occurs after clearance of the lesion scar, Signal Transduct. Targeted Ther. 7 (1) (2022) 184.
[16] Y. Jing, et al., Fecal microbiota transplantation exerts neuroprotective effects in a mouse spinal cord injury model by modulating the microenvironment at the lesion site, Microbiol. Spectr. 10 (3) (2022), e0017722.
[17] J. Gao, et al., Lipoic acid enhances survival of transplanted neural stem cells by reducing transplantation-associated injury, J. Neurorestoratol. 1 (2013).
[18] T.T. Roberts, G.R. Leonard, D.J. Cepela, Classifications in brief: American spinal injury association (ASIA) impairment scale, Clin. Orthop. Relat. Res. 475 (5) (2017) 1499–1504.
[19] T. Liebscher, et al., Cervical spine injuries with acute traumatic spinal cord injury: spinal surgery adverse events and their association with neurological and functional outcome, Spine (Phila Pa 1976) 47 (1) (2022) E16–E26.
[20] B.K. Kwon, et al., Neurochemical biomarkers in spinal cord injury, Spinal Cord 57 (10) (2019) 819–831.
[21] K. Fouad, et al., The neuroanatomical-functional paradox in spinal cord injury, Nat. Rev. Neurol. 17 (1) (2021) 53–62.
[22] M. Shu, et al., Single-cell RNA sequencing reveals Nestin(+) active neural stem cells outside the central canal after spinal cord injury, Sci. China Life Sci. 65 (2) (2022) 295–308.
[23] L.M. Milich, et al., Single-cell analysis of the cellular heterogeneity and interactions in the injured mouse spinal cord, J. Exp. Med. 218 (8) (2021).
[24] B. Wang, et al., Spinal cord injury target-immunotherapy with TNF-alpha autoregulated and feedback-controlled human umbilical cord mesenchymal stem cell derived exosomes remodelled by CRISPR/Cas9 plasmid, Biomater. Adv. 133 (2022), 112624.
[25] A. Cugurra, et al., Skull and vertebral bone marrow are myeloid cell reservoirs for the meninges and CNS parenchyma, Science 373 (6553) (2021).

[26] R.N. Beladi, et al., Serine proteases and chemokines in neurotrauma: new targets for immune modulating therapeutics in spinal cord injury, Curr. Neuropharmacol. 19 (11) (2021) 1835–1854.
[27] J. Zhao, et al., Emerging role of PD-1 in the central nervous system and brain diseases, Neurosci. Bull. 37 (8) (2021) 1188–1202.
[28] S. Sharif, M.Y. Jazaib Ali, Outcome prediction in spinal cord injury: myth or reality, World Neurosurg 140 (2020) 574–590.
[29] L.D. Hachem, C.S. Ahuja, M.G. Fehlings, Assessment and management of acute spinal cord injury: from point of injury to rehabilitation, J Spinal Cord Med 40 (6) (2017) 665–675.
[30] T. Jogia, et al., Peripheral white blood cell responses as emerging biomarkers for patient stratification and prognosis in acute spinal cord injury, Curr. Opin. Neurol. 34 (6) (2021) 796–803.
[31] N. Kyritsis, et al., Diagnostic blood RNA profiles for human acute spinal cord injury, J. Exp. Med. 218 (3) (2021).
[32] A. Abdelhak, et al., Blood GFAP as an emerging biomarker in brain and spinal cord disorders, Nat. Rev. Neurol. 18 (3) (2022) 158–172.
[33] J.W. Hammond, et al., Complement-dependent synapse loss and microgliosis in a mouse model of multiple sclerosis, Brain Behav. Immun. 87 (2020) 739–750.
[34] G.T. Norris, et al., Neuronal integrity and complement control synaptic material clearance by microglia after CNS injury, J. Exp. Med. 215 (7) (2018) 1789–1801.
[35] I. Michailidou, et al., Complement C1q-C3-associated synaptic changes in multiple sclerosis hippocampus, Ann. Neurol. 77 (6) (2015) 1007–1026.
[36] A.N. Stewart, et al., Acute inflammatory profiles differ with sex and age after spinal cord injury, J. Neuroinflammation 18 (1) (2021) 113.
[37] K.R. Byrnes, et al., Delayed inflammatory mRNA and protein expression after spinal cord injury, J. Neuroinflammation 8 (2011) 130.
[38] Z. Ren, et al., Gal-3 is a potential biomarker for spinal cord injury and Gal-3 deficiency attenuates neuroinflammation through ROS/TXNIP/NLRP3 signaling pathway, Biosci. Rep. 39 (12) (2019).
[39] I. Yanatori, et al., CD63 is regulated by iron via the IRE-IRP system and is important for ferritin secretion by extracellular vesicles, Blood 138 (16) (2021) 1490–1503.
[40] N.Z. Khan, et al., Spinal cord injury alters microRNA and CD81+ exosome levels in plasma extracellular nanoparticles with neuroinflammatory potential, Brain Behav. Immun. 92 (2021) 165–183.
[41] S. Salvany, et al., Microglial recruitment and mechanisms involved in the disruption of afferent synaptic terminals on spinal cord motor neurons after acute peripheral nerve injury, Glia 69 (5) (2021) 1216–1240.