Pharmacogenomics and Personalized Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Bacterial Lipopolysaccharide-Associated Genes and Molecular Subtypes in Autism Spectrum Disorder

Yuanxia He[1,2,*], Yun He[2,*], Boli Cheng[1,2]

[1]Department of Clinical Medicine, North Sichuan Medical College, Nanchong, Sichuan, 637000, People's Republic of China; [2]Department of Pediatrics, Affiliated Hospital, North Sichuan Medical College, Nanchong, Sichuan, 637000, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yun He; Boli Cheng, Department of Pediatrics, Affiliated Hospital, North Sichuan Medical College, Nanchong, 637000, Sichuan, People's Republic of China, Email heyun02409@aliyun.com

Background: Autism spectrum disorder (ASD) is a complex neurodevelopmental condition marked by diverse symptoms affecting social interaction, communication, and behavior. This research aims to explore bacterial lipopolysaccharide (LPS)- and immune-related (BLI) molecular subgroups in ASD to enhance understanding of the disorder.

Methods: We analyzed 89 control samples and 157 ASD samples from the GEO database, identifying BLI signatures using least absolute shrinkage and selection operator regression (LASSO) and logistic regression machine learning algorithms. A nomogram prediction model was developed based on these signatures, and we performed Gene Set Enrichment Analysis (GSEA), Gene Set Variation Analysis (GSVA), and immune cell infiltration analysis to assess the impact of BLI subtypes and their underlying mechanisms.

Results: Our findings revealed 17 differentially expressed BLI genes in children with ASD, with BLNK, MAPK8, PRKCQ, and TNFSF12 identified as potential biomarkers. The nomogram demonstrated high diagnostic accuracy for ASD. We delineated two distinct molecular subtypes (Cluster 1 and Cluster 2), with GSVA indicating that Cluster 2 showed upregulation of immune- and inflammation-related pathways. This cluster exhibited increased levels of antimicrobial agents, chemokines, cytokines, and TNF family cytokines, alongside activation of bacterial lipoprotein-related pathways. A significant correlation was found between these pathways and distinct immune cell subtypes, suggesting a potential mechanism for neuroinflammation and immune cell infiltration in ASD.

Conclusion: Our research highlights the role of BLI-associated genes in the immune responses of individuals with ASD, indicating their contribution to the disorder's typification. The interplay between bacterial components, genetic predisposition, and immune dysregulation offers new insights for understanding ASD and developing personalized interventions.

Keywords: autism, bacterial metabolites, immune cell infiltration, GEO, molecular subtypes, immune responses

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition that impacts about 2.3% of 8-year-old children in the United States, manifesting through challenges in social communication and the display of limited, repetitive behaviors or interests.[1,2] The heterogeneity of its symptoms and the wide-ranging severity make ASD a particularly challenging disorder to understand and manage.[3,4] The etiology of ASD is multifactorial, involving a combination of genetic, environmental, and neurobiological factors.[5] Among these, the dysregulation of the immune system and the impact of bacterial metabolites have emerged as critical areas of interest, shedding light on novel pathophysiological pathways and potential therapeutic targets.[6,7] It is important to clarify that ASD is not an infectious condition.

Recent studies have highlighted the role of the immune system in the development and progression of ASD, suggesting that immune dysregulation may contribute to the neuroinflammatory processes observed in individuals with the disorder.[8,9] Concurrently, the gut-brain axis and its microbial constituents have been implicated in ASD, where alterations in the gut

Received: 1 September 2024
Accepted: 31 December 2024
Published: 17 January 2025

microbiome may influence neurodevelopment and behavior through the production of bacterial metabolites, including lipopolysaccharides (LPS).[10,11] LPS, a component of the outer membrane of gram-negative bacteria, is known to trigger robust immune responses and may play a role in the neuroinflammatory pathways associated with ASD.[12] Additionally, maternal immune activation during pregnancy has been increasingly recognized as a significant environmental factor in the etiology of ASD. Maternal immune activation, often induced by infections or other immune challenges during pregnancy, can lead to the activation of the maternal immune system and subsequent inflammatory responses that may affect fetal brain development, contributing to the development of ASD in offspring.[13,14]

The primary objective of this research is to explore the intersection of bacterial LPS and immune response-related gene signatures in the classification of ASD. By identifying specific molecular subgroups associated with these factors, we aim to deepen our understanding of the biological underpinnings of ASD and pave the way for more targeted diagnostic and therapeutic strategies. Methods leveraging high-throughput genomic data and sophisticated machine learning algorithms, such as the least absolute shrinkage and selection operator (LASSO) regression and logistic regression, have proven invaluable in dissecting the complex genetic landscape of diseases like ASD. These approaches enable the identification of biomarkers and molecular signatures with high precision, offering insights into disease mechanisms and potential intervention points.[15,16]

Our study utilizes data from the GSE18123 and GSE42133 datasets, employing advanced bioinformatics analyses, including Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA), to unravel the biological processes and pathways associated with the identified bacterial lipopolysaccharide and immune-related (BLI) gene signatures in ASD. These analyses provide a comprehensive view of the functional implications of the BLI signatures, shedding light on the immune and inflammatory pathways that may be dysregulated in ASD. The identification of distinct molecular subtypes based on BLI signatures holds the promise of enhancing our comprehension of ASD's heterogeneity. By delineating the molecular underpinnings associated with different ASD phenotypes, we can move towards a more nuanced understanding of the disorder, facilitating the development of personalized interventions tailored to the specific biological profiles of individuals with ASD. The diagram in Figure 1 illustrated the process of the current research. In summary, this study highlights the importance of BLI-related genes in the immune responses of individuals with ASD. The research indicates that these genes are key in characterizing the disorder, providing fresh insights into the link between bacterial elements, genetic predisposition, and immune dysregulation. This research paves the way for understanding the complexities of ASD and emphasizes the potential for tailored therapeutic interventions based on molecular subtypes.

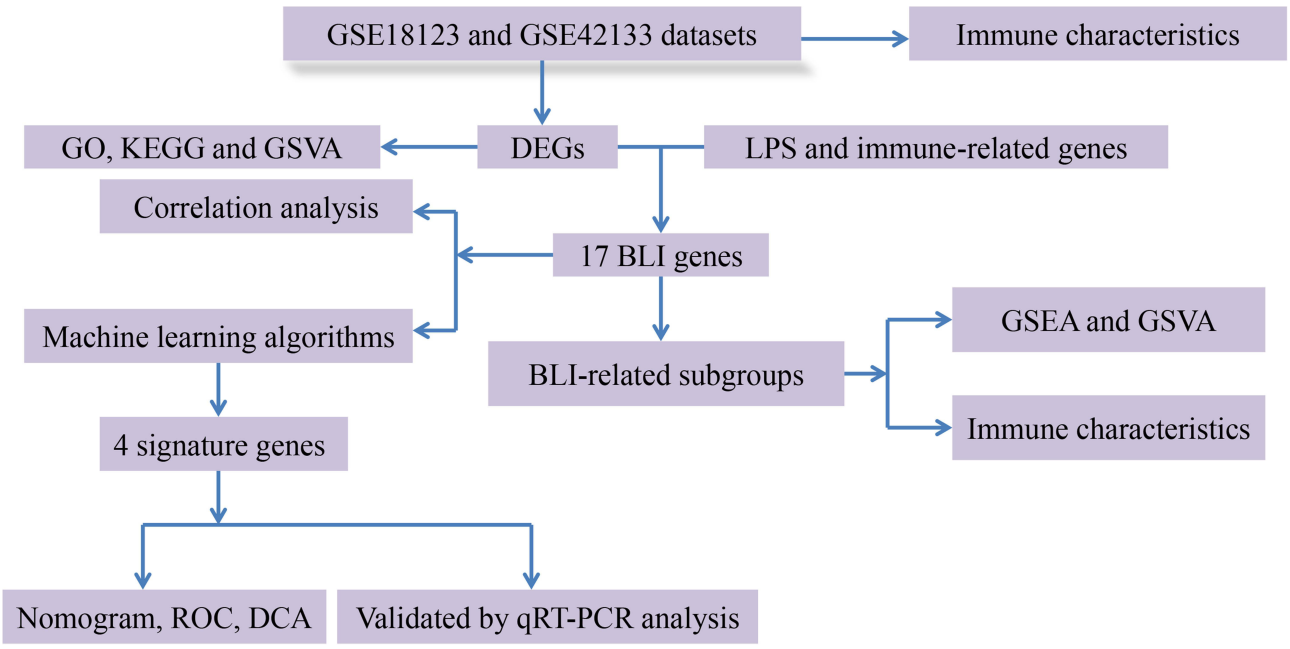

Figure 1 Visual representation of the research process.

Methods

Acquisition and Initial Processing of Data

Two sets of genetic expression microarrays associated with ASD (GSE18123 and GSE42133) were acquired from the Gene Expression Omnibus (GEO) database. Details of the datasets can be found in Table 1. The GSE18123 dataset included 33 samples from control subjects and 66 samples from individuals with ASD, while the GSE42133 dataset comprised 56 control samples and 91 ASD samples. The GEO datasets were prepared and standardized through the utilization of the affy R package. Subsequently, the two datasets were combined with the assistance of the sva package in R to mitigate batch discrepancies. Following this amalgamation, the samples from the combined dataset were separated into control (n = 89) and ASD (n = 157) groups. It is important to note that the datasets do not provide information regarding whether the ASD patients had any known bacterial infections prior to their diagnosis. We utilized the Comparative Toxicogenomics Database (CTD, http://ctdbase.org/) to collect the lipopolysaccharide-related genes (LRGs) and obtained the immune-related genes (IRGs) from the Immunology Database and Analysis Portal (ImmPort) (https://www.immport.org/shared/home).

Analysis of Differentially Expressed Genes (DEGs) and Functional Enrichment Analysis

We employed the limma package to identify DEGs between the control and ASD groups. Genes meeting the criteria for differential expression were those with a p-value adjusted to below 0.05 and a |logFC| greater than 0.2. Subsequently, we utilized the ggplot2 and ComplexHeatmap packages to generate the volcano plot and heatmap, respectively. For GO and KEGG enrichment analysis, we utilized the clusterProfiler and enrichplot packages. A statistically significant threshold was set at an adjusted p-value less than 0.05.

Analysis of the Immune Microenvironment

We undertook a comprehensive analysis of immunological profiles, encompassing various immune cell subsets and the immune reaction pathways. The immune cell subpopulations were quantified employing the xCell tool and single-sample Gene Set Enrichment Analysis (ssGSEA). Immune-related gene constellations were acquired from the ImmPort repository (http://www.immport.org) for the purpose of quantifying the activity of 16 immune reaction pathways through ssGSEA. Visualization of the findings was facilitated by the ggplot2 package. The Pearson's correlation coefficient was calculated to elucidate the correlations between the bacterial lipoprotein-associated pathways and the proportions of immune cells, as well as the immune reaction pathways. Statistical significance was established at a P-value threshold of <0.05.

Discovery of Signature Genes

The BLI genes were obtained through the intersection of genes derived from DEGs, LRGs, and IRGs. Utilizing the glmnet package, a LASSO analysis was conducted on the expression profiles of BLI genes, incorporating penalty parameters for 10-fold cross-validation. Subsequently, a logistic regression analysis was employed to identify signature genes.

Development and Validation of the Nomogram

To construct the nomogram, we utilized the rms R package, integrating the signature genes. The nomogram's ability to diagnosis ASD was evaluated by creating a Receiver Operating Characteristic (ROC) curve. Furthermore, the accuracy of the nomogram was assessed using calibration curves and decision curve analysis (DCA).

Table 1 Dataset Information Related to ASD in the GEO Database

GEO ID	Platform	Control group	Disease group	Source
GSE18123	GPL570	33	66	Blood
GSE42133	GPL10558	56	91	Blood

https://doi.org/10.2147/PGPM.S494126

Identifying Subcategories of Diseases Associated with BLI

To explore the connection between BLI and ASD, we utilized the ConsensusClusterPlus tool for cluster analysis on BLI expression profiles. The ideal cluster count, labeled as "k" value, was determined from the Cumulative Distribution Function (CDF) graph. Additionally, we performed variance analysis on different disease subcategories using the limma package.

Gene Set Enrichment Analysis (GSEA)

GSEA is a computational method utilized to assess whether a predefined set of genes exhibits statistically significant, concordant differences between two biological states.[17] Utilizing the GSEA tool, we categorized genes based on their differential expression levels among distinct subgroups. A specific dataset named "h.all.v7.4.symbols.gmt" was sourced from the Molecular Signatures Database. Each subgroup was subjected to assessment involving 1000 permutations to compute enrichment scores (ES) for significance determination, and to establish statistical significance using p-values and false discovery rate (FDR). Gene sets were considered significant if they displayed p-values below 0.05 and FDR values below 0.25.

Gene Set Variation Analysis (GSVA)

To explore the differences in biological pathways, we utilized the Gene Set Variation Analysis (GSVA) approach.[18] The GSVA was performed through the tidyverse package, with the "c5.all.v7.1.symbols.gmt" dataset being sourced as the reference gene collection. Subsequently, the differential pathways in the Cluster 1 and Cluster 2 subgroups were identified using the limma package. The criteria for conducting the differential expression examination were set at p.adj < 0.05 and log Fold Change (FC) ≥ 0.2. The visualization of the outcomes from the volcano plot and heatmap was achieved using the "ggplot2" and "ComplexHeatmap" packages.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR) Analysis

Between the years 2022 and 2023, a total of 8 individuals diagnosed with ASD and 8 healthy individuals (HC) were recruited from the North Sichuan Medical College to participate in a study involving the collection of blood specimens. The blood samples underwent RNA extraction utilizing TRIzol Reagent (ThermoFisher Scientific, USA) in accordance with the provided guidelines to acquire total RNA. The isolated RNA was then employed for cDNA synthesis with the aid of cDNA synthesis kits (ThermoFisher Scientific, USA). Subsequent to this, the qRT-PCR analysis was conducted utilizing the ABI 7900HT PCR system (Applied Biosystems, USA). The relative levels of mRNA expression were determined utilizing the $2^{-\Delta\Delta Ct}$ methodology. Primer sequences are shown in Table S1.

Results

Differential Gene Expression and Pathway Enrichment Analysis Between Control and ASD Groups

As shown in Figure 2A, volcano plot analysis indicated that 377 DEGs were significantly upregulated (red dots) and 118 DEGs were significantly downregulated (blue dots) in the ASD group compared to the control group. The hierarchical clustering heatmap (Figure 2B) provided a visual representation of these DEGs, with clear clustering patterns differentiating the ASD samples from the control samples. Functional enrichment analysis (Figure 2C) showed that these DEGs were significantly enriched in biological pathways related to lipid and atherosclerosis, osteoclast differentiation, and MAPK signaling pathways, among others. Notable are the processes involved in immune response such as Fc receptor signaling pathways, acute myeloid leukemia, and positive regulation of leukocyte activation, which could hint at an immune component in the etiology of ASD. GSVA results (Figure 2D) provided further insight into the functional implications of the altered gene expression. Several pathways showed significant differences between the two groups, with the adipocytokine and neurotrophin signaling pathways being highly enriched in the ASD group. Importantly, pathways associated with metabolism, signaling, and neurodevelopment were disrupted, suggesting a multifactorial impact on cellular and systemic functions that may contribute to the pathology of ASD. In summary, these analyses

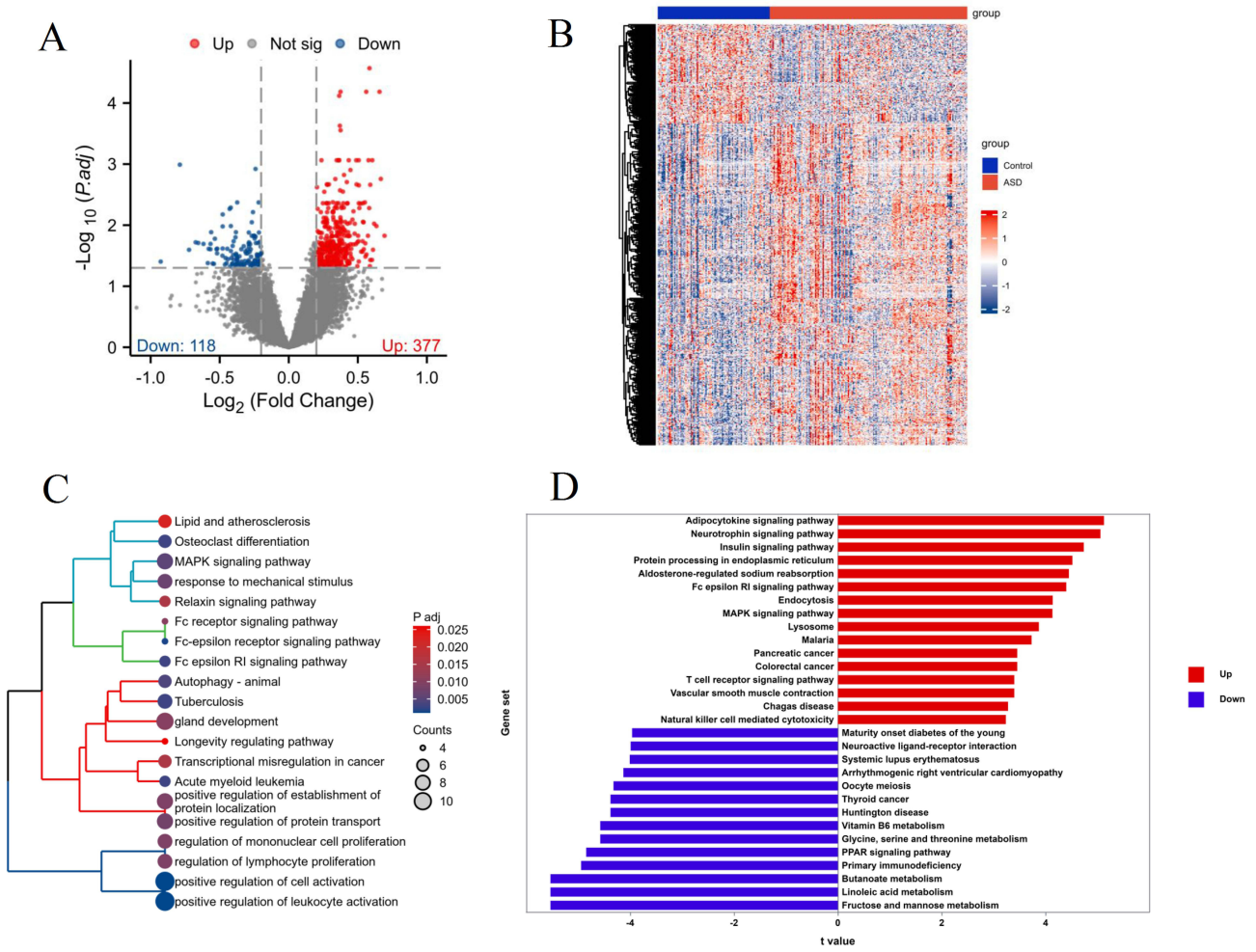

Figure 2 Differential gene expression and pathway enrichment analyses in ASD. (**A**) The volcano plot depicts the differential gene expression between control and ASD groups, with red dots signifying upregulated genes and blue dots indicating downregulated genes. (**B**) The heatmap illustrates the results of DEGs based on their expression levels in control and ASD samples. Each row represents an individual gene, and each column represents a sample. The color gradient, from blue to red, reflects gene expression levels, spanning from downregulated to upregulated. (**C**) Functional enrichment analysis of DEGs. The size of each bubble corresponds to the number of DEGs associated with each term, and the colors represent the range of adjusted p-values, with darker colors indicating more significant enrichment. (**D**) The bar chart shows the results of GSVA for DEGs between control and ASD groups. Red bars indicate pathways with a significant enrichment of upregulated genes, blue bars represent pathways with downregulated genes, and the length of each bar corresponds to the t-value, signifying the degree of enrichment.

delineated patterns of gene expression and molecular pathways that differentiate the ASD group from controls, providing a foundation for further investigation into the molecular underpinnings of ASD.

Identification and Correlation Analysis of BLI in ASD

Venn diagram (Figure 3A) illustrates the intersections between DEGs, LRGs, and IRGs in the context of ASD. The analysis identified 17 BLI at the intersection of all three categories, representing potential key players in the molecular pathways involved in ASD that are responsive to bacterial endotoxins and have a role in immunity. The network analysis (Figure 3B) highlights the intricate correlation patterns among the BLI genes identified. The nodes represent individual genes, while the edges between them indicate the strength of the expression correlation. The network showcases several genes with high connectivity, suggesting these may serve as central hubs in the regulatory network affected by ASD. The heatmap (Figure 3C) shows the expression analysis of the identified BLI genes in the control and ASD groups. Hierarchical clustering reveals distinct patterns of gene expression, with marked differential expression between the two groups. In summary, this analysis provides insights into the genes implicated in the response to bacterial endotoxins

https://doi.org/10.2147/PGPM.S494126

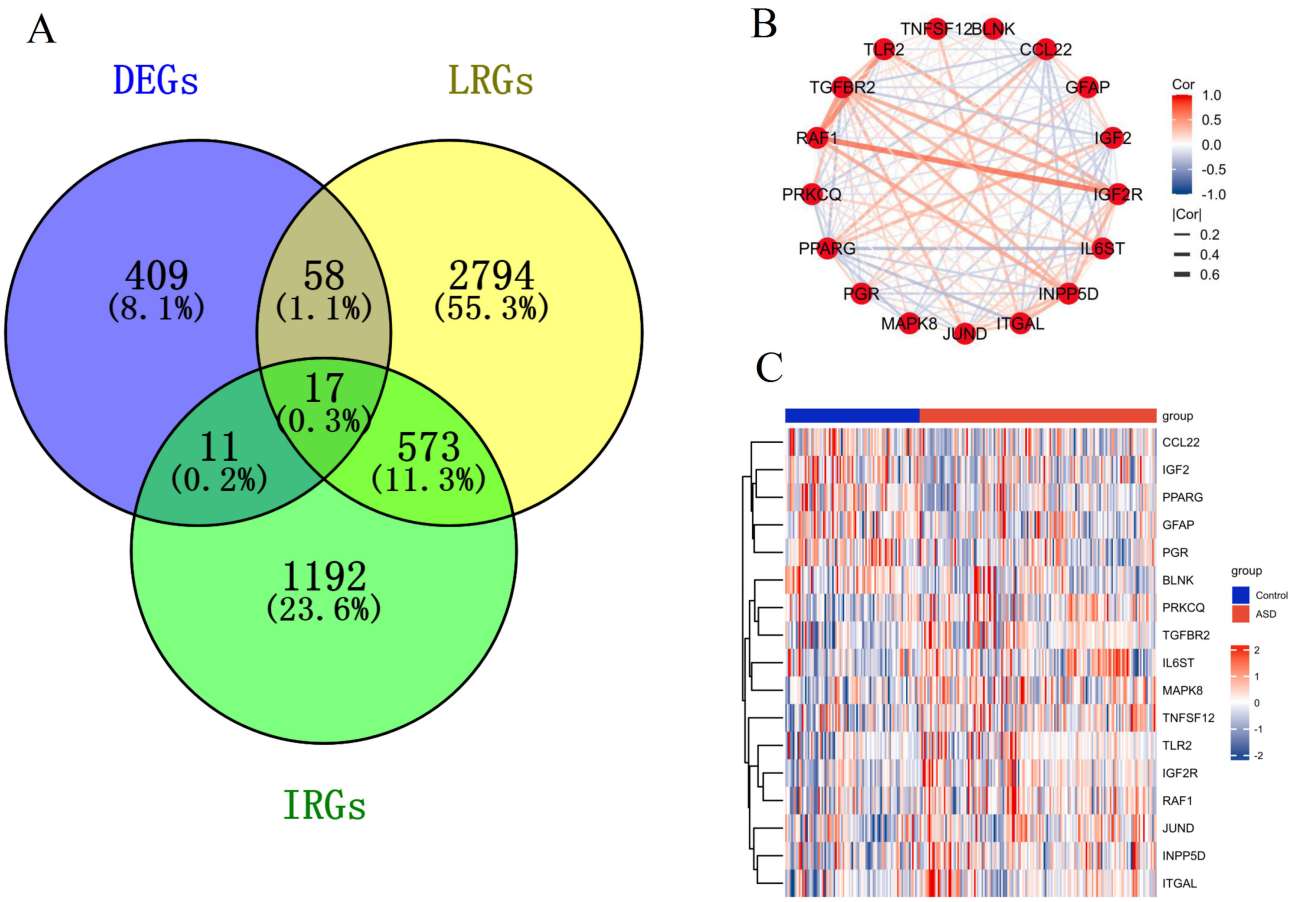

Figure 3 Integrative analysis of BLI genes in ASD. (**A**) The Venn diagram displays the overlaps among DEGs (blue circle), LRGs (yellow circle), and IRGs (green circle). The numbers in each compartment indicate the quantity of genes unique to or shared between the categories. (**B**) Correlation network of BLI genes. Nodes represent BLI genes, and edges show the correlation between gene expression levels, with red indicating positive and blue indicating negative correlations. (**C**) Heatmap of BLI gene expression. Rows correspond to individual genes, and columns represent samples from each group. The color gradient from blue (low expression) to red (high expression) illustrates the gene expression patterns.

and immune processes in ASD. The identified BLI genes and their intricate correlation network could serve as a foundation for further exploration of their roles in the pathogenesis and progression of ASD.

Assessment of Immune Cell Infiltration and Immune-Related Pathway Levels

The ssGSEA analysis (Figure 4A) revealed distinct immune cell infiltration patterns between the control and ASD groups. Notably, several immune cell types, such as B cells, DC, iDC, and Th1 cells, showed significant differences in infiltration levels, with an overall decrease observed in the ASD group. The xCell scores (Figure 4B) provide further detail on the immune cell composition within the control and ASD samples. Specific cell populations, including aDC, B cells, memory B cells, naïve B cells, NKT, pro B cells, and Th1 cells, exhibited reduced levels in the ASD group. The results suggest a divergence in the immune landscape, potentially contributing to the pathophysiology of ASD. Analysis of immune-related pathway activity (Figure 4C) using ssGSEA demonstrated altered pathway engagement between the two groups. The levels of chemokines, cytokines, interleukins, and TGFb family member were markedly lower in the ASD group. These pathways are critical for robust immune responses, and their reduced activity suggests potential dysregulation of immune function in ASD. Together, these findings underline notable variances in the immune cell milieu and pathway activation between ASD and control samples, providing valuable insights into the altered immunological states associated with ASD.

https://doi.org/10.2147/PGPM.S494126

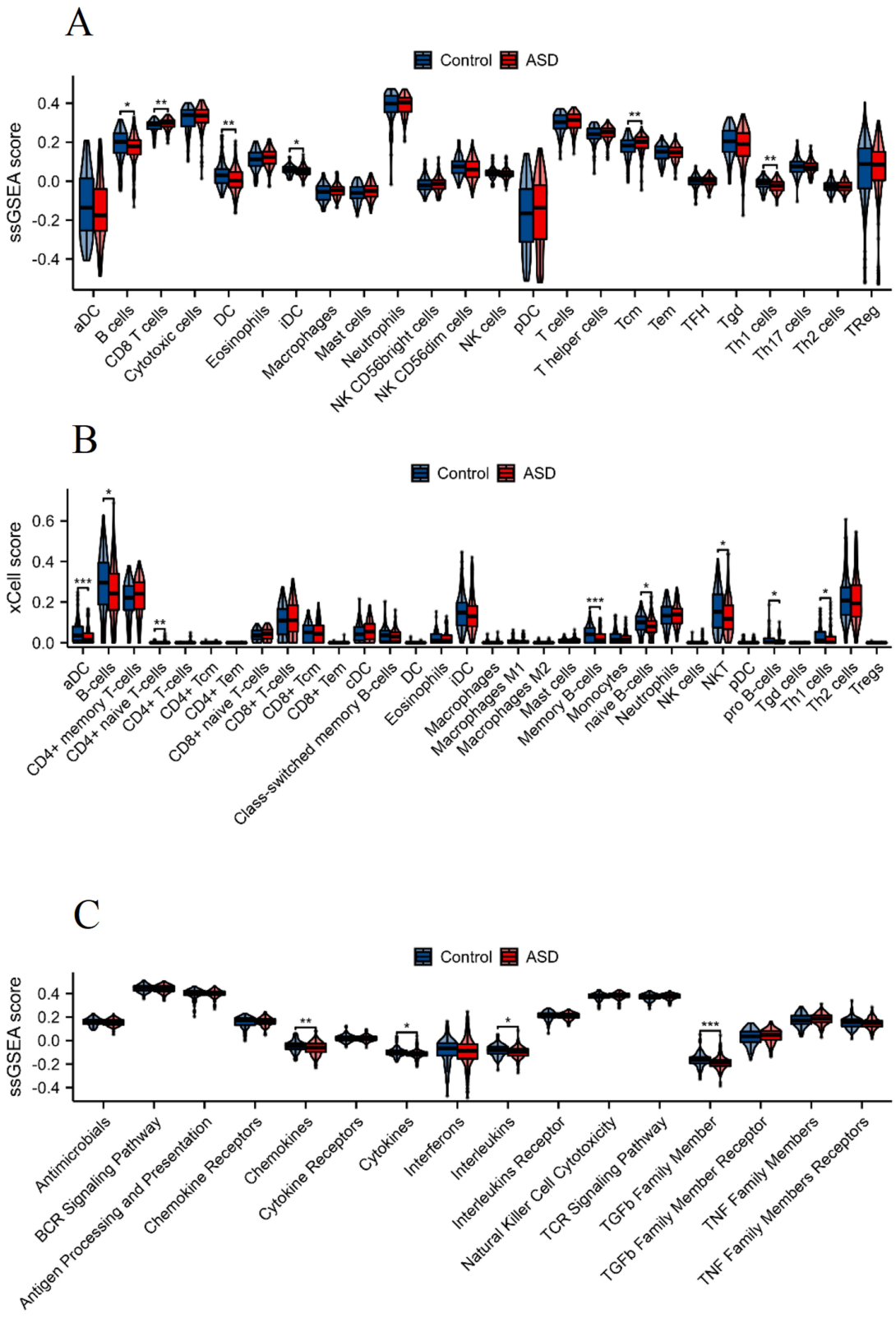

Figure 4 Evaluation of immune cell infiltration and pathway activation differences in control and ASD groups. The violin plots illustrate the distribution of ssGSEA scores (**A**) and xCell scores (**B**) for various immune cell types in the control (blue) and ASD (red) groups, indicating the relative level of immune infiltration for each cell type in the respective cohorts. (**C**) The violin plots display the ssGSEA pathway scores for immune-related pathways in the control (blue) and ASD (red) groups. *p < 0.05, **p < 0.01, ***p < 0.001.

https://doi.org/10.2147/PGPM.S494126

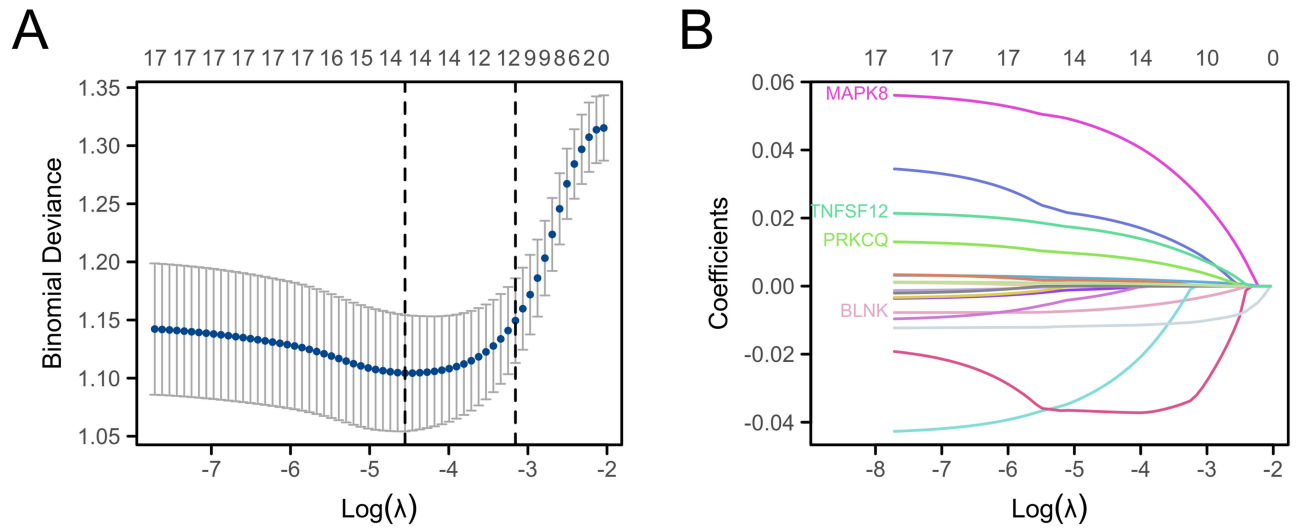

Figure 5 LASSO regression analysis identifies key BLI genes in ASD. (**A**) The model selection process in the LASSO regression analysis, with the y-axis displaying the binomial deviance and the x-axis showing the log-transformed regularization parameter (log(λ)). The blue dots are the average deviance values from cross-validation, and the vertical bars represent the standard error. (**B**) The coefficient profiles of genes are plotted against different log(λ) values in the LASSO regression analysis. Each colored line represents the coefficient of a gene. Genes with coefficients shrinking to zero under increasing regularization are filtered out, leaving those with non-zero coefficients, which are considered as key contributors to the model.

Identification of Key BLI Genes in ASD Using LASSO Regression and Logistic Regression

As shown in Figure 5A and B, the LASSO regression technique was conducted to identify biomarkers that could serve as indicators for ASD. Our results highlighted four genes - BLNK, MAPK8, PRKCQ, TNFSF12 - as having a substantial association with ASD, as evidenced by their p values falling below the threshold of 0.05 in logistic regression evaluations, detailed in Table 2.

Expression Analysis and Construction of a Nomogram

The ASD group exhibited decreased expression of the BLNK gene compared to the control group (Figure 6A). Conversely, the genes MAPK8 (Figure 6B), PRKCQ (Figure 6C), and TNFSF12 (Figure 6D) showed elevated expression levels in the ASD group relative to the control group. A nomogram was constructed (Figure 6E) using the expression levels of these key BLI genes as predictive variables to estimate the probability of ASD. This nomogram integrates the individual scores for BLNK, MAPK8, PRKCQ, and TNFSF12 to generate a total score, which correlates with the estimated probability of ASD. The nomogram's predictive performance was quantified by the area under the receiver operating characteristic curve (AUC) (Figure 6F). The model displayed good discriminatory ability with an AUC of 0.761, and confidence intervals (CI) suggesting the model's robustness. Calibration curves (Figure 6G) indicate the agreement between the predicted and actual probabilities of ASD, with the red line representing model performance. The closer the red line is to the ideal grey line, the better the predictions align with actual outcomes. Decision curve analysis

Table 2 The Results of Logistic Regression

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Odds Ratio (95% CI)	P value	Odds Ratio (95% CI)	P value
BLNK	246	1.004 (1.001–1.007)	**0.005**	1.008 (1.004–1.012)	**< 0.001**
MAPK8	246	0.941 (0.910–0.972)	**< 0.001**	0.947 (0.910–0.985)	**0.007**
PRKCQ	246	0.992 (0.987–0.997)	**0.003**	0.989 (0.982–0.996)	**0.001**
TNFSF12	246	0.975 (0.960–0.990)	**0.002**	0.981 (0.962–0.999)	**0.039**

Note: Bold values indicate statistically significant differences.

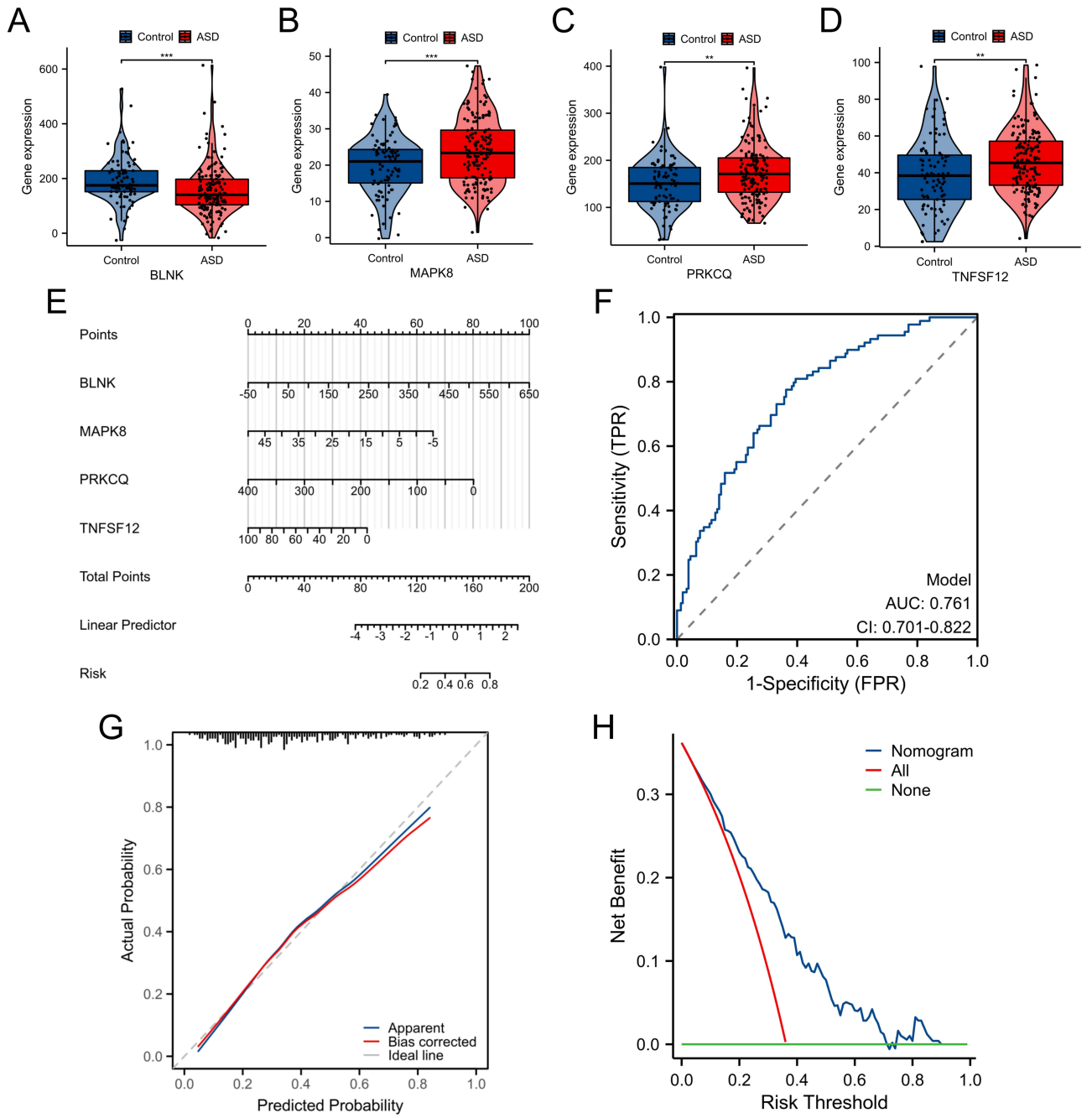

Figure 6 Evaluating the expression levels of key BLI genes and constructing a predictive nomogram for ASD. This violin plot displays the expression levels of BLNK (**A**), MAPK8 (**B**), PRKCQ (**C**), and TNFSF12 (**D**) in the control (blue) and ASD (red) groups. **p < 0.01, ***p < 0.001. (**E**) The nomogram is a graphical representation designed to predict the likelihood of ASD by integrating the expression values of BLNK, MAPK8, PRKCQ, and TNFSF12. (**F**) The ROC curve demonstrates the predictive accuracy of the ASD nomogram. (**G**) This calibration plot assesses the agreement between predicted probabilities of ASD and observed outcomes. The red line illustrates the performance of the nomogram, with the dashed grey line representing the ideal scenario where predictions perfectly match observed events. (**H**) The decision curve quantifies the clinical net benefit of using the nomogram across different decision thresholds.

(Figure 6H) evaluates the clinical usefulness of the nomogram by calculating the net benefit across various risk thresholds. The nomogram (blue line) demonstrates superior benefit in comparison to the treat-all (red line) or treat-none (green line) strategies over a range of risk thresholds, supporting the nomogram's practical value in clinical decision-making. In summary, these results suggest that the identified key BLI genes are differentially expressed in ASD and that the developed nomogram can effectively estimate the likelihood of ASD, potentially assisting in early diagnosis and intervention strategies.

https://doi.org/10.2147/PGPM.S494126

Unsupervised Clustering Analysis of ASD Patients

The consensus matrix from the unsupervised analysis (Figure 7A) displays the clustering of ASD patient samples into two primary groups, Cluster 1 (C1) and Cluster 2 (C2). The elbow plot (Figure 7B) was utilized to identify the most appropriate number of clusters (K). The curve's elbow point, which corresponds to a sharp change in the descent slope, indicates an optimal cluster number of K = 2 for this dataset. The CDF (Figure 7C) further validates the choice of K by comparing the consensus distribution for different numbers of K. The flattening of the CDF curves at K = 2 reinforces the elbow plot's indication that two clusters provide a stable and reliable grouping of the ASD samples based on the expression of the BLI genes. The heatmap (Figure 7D) visualizes the gene expression patterns of the 17 BLI genes across the two identified clusters. The color gradient from blue to red indicates expression levels from low to high, respectively. The hierarchical clustering dendrograms on the top and left sides of the heatmap showcase the grouping of both samples and genes into clusters. This stratification reflects the potentially distinct molecular subtypes of ASD defined by

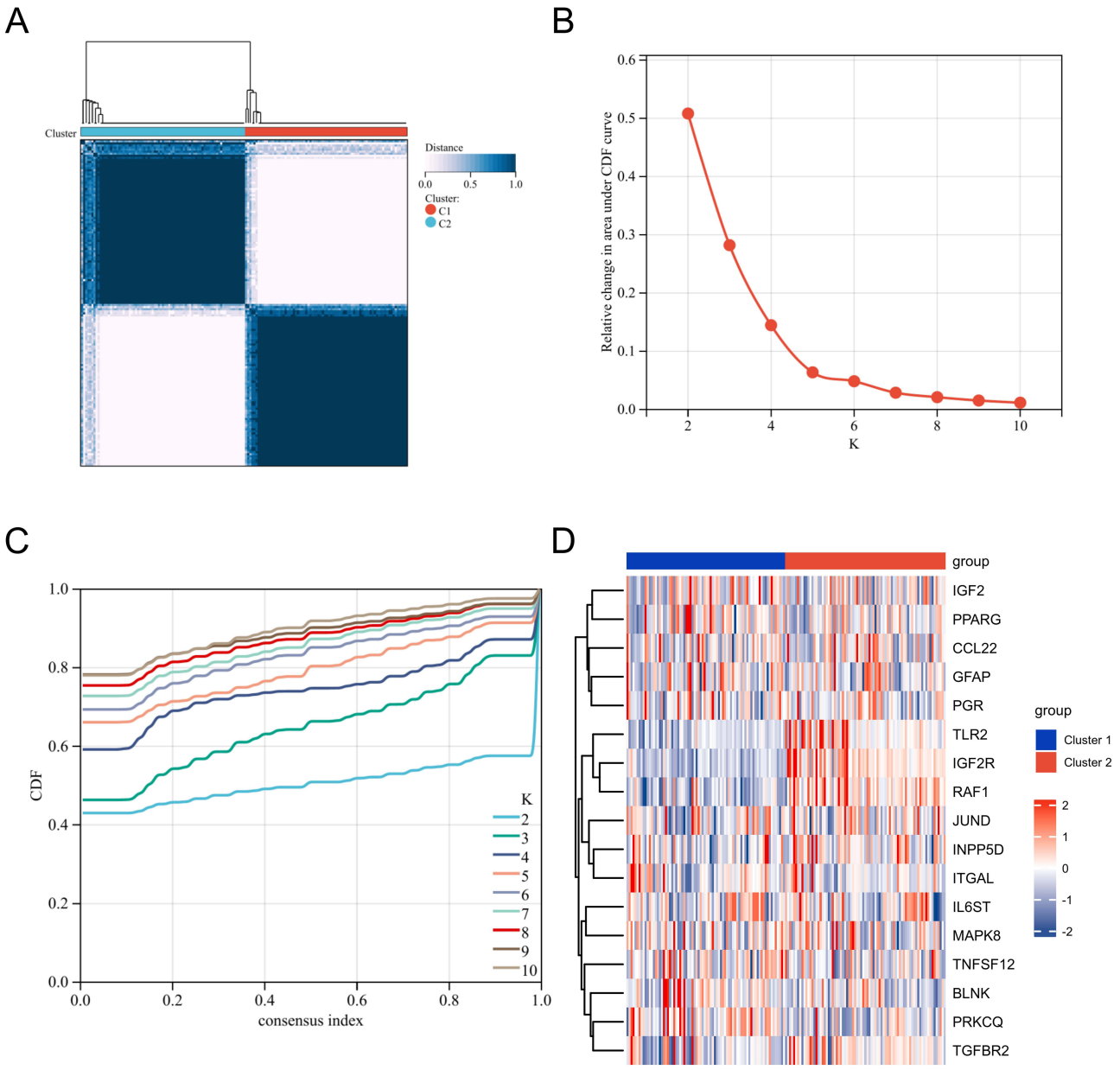

Figure 7 Stratification of ASD patients into distinct clusters. (**A**) Consensus clustering matrix. (**B**) Elbow plot for determining optimal clusters. (**C**) CDF for cluster validation. (**D**) Heatmap of 17 BLI gene expression in ASD clusters.

https://doi.org/10.2147/PGPM.S494126

differential expression of the BLI genes. In conclusion, the unsupervised clustering based on the BLI gene expression profiles effectively divided ASD patients into two distinct clusters, suggesting the presence of molecularly distinct subtypes within the ASD population.

Differential Expression and Enrichment Analysis of BLI-Related Subgroups in ASD

The volcano plot (Figure 8A) displays the differential expression analysis results between Cluster 1 and Cluster 2. In total, 1,583 genes were significantly upregulated (red dots) and 824 genes were significantly downregulated (blue dots) in Cluster 1 compared to Cluster 2. Figure 8B presents a heatmap depicting the standardized expression levels of differentially expressed genes across the two clusters. The color gradient from blue to red signifies the expression levels from low to high, respectively. Hierarchical clustering was applied to group genes with similar expression patterns, differentiating Cluster 1 from Cluster 2. GSEA was employed to investigate the enrichment of biological pathways in the differentially expressed genes. The pathways related to immune function, cytokine signaling, inflammasomes, and natural killer cell-mediated cytotoxicity are particularly pronounced, indicating distinct immunological profiles in the different clusters (Figure 8C). In Figure 8D, a different volcano plot is presented to identify the distinct pathways between Cluster 1 and Cluster 2 by. A total of 725 pathways showed significant upregulation (represented by red dots), while 248 pathways exhibited significant downregulation (represented by blue dots) in Cluster 1 compared to Cluster 2. The Figure 8E heatmap illustrates the GSVA outcomes, showing a significant enrichment of inflammatory and immune-related pathways in the Cluster 2 group compared to Cluster 1. In conclusion, the differential expression and enrichment analysis revealed significant variations

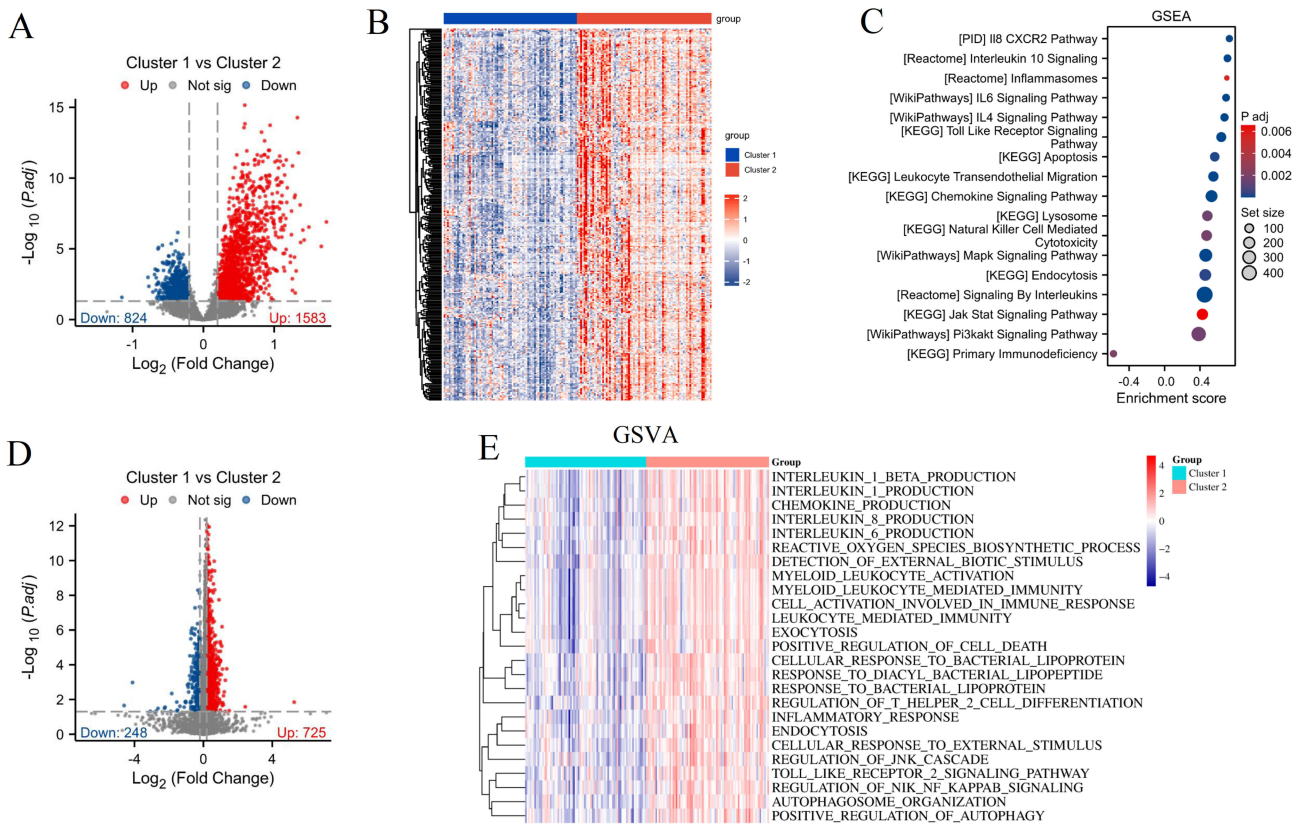

Figure 8 Differential gene expression and pathway enrichment analysis in ASD. (**A**) Volcano plot illustrates the differential gene expression between Cluster 1 and Cluster 2, where red dots depict significantly upregulated genes, blue dots show significantly downregulated genes, and gray dots represent genes with no significant change. (**B**) The heatmap visualizes the differences in gene expression between Cluster 1 and Cluster 2. Red indicates higher expression, blue signifies lower expression. (**C**) The GSEA results show significant biological pathways between Clusters. (**D**) Volcano plot illustrates the differential pathways between Cluster 1 and Cluster 2, where red dots depict significantly upregulated pathways, blue dots show significantly downregulated pathways, and gray dots represent pathways with no significant change. (**E**) The GSVA heatmap differentiates the activity of biological processes between Cluster 1 and Cluster 2. Colored from red to blue, red denotes higher activity and blue lower activity of gene sets.

https://doi.org/10.2147/PGPM.S494126

between ASD subgroups that are BLI-related. This underscores potential differences in immune responses and inflammation that may be relevant for subgroup-specific therapeutic approaches or prognostic assessments.

Comparative Assessment of Immune Cell Infiltration and Pathway Activation in ASD Subgroups

In this study, we utilized ssGSEA and xCell algorithms to assess immune cell infiltration and activity of immune-related pathways in subgroups related to BLI in patients with ASD. The ssGSEA-derived immune cell infiltration scores for various immune cell types in two distinct ASD clusters, known as Cluster 1 and Cluster 2, are illustrated in Figure 9A. Violin plots reveal notable differences in the abundance of immune cells, with Cluster 1 showing elevated levels of infiltration across B cells, T cells, T helper cells, TFH, Th1 cells, and Th2 cells, while displaying reduced infiltration levels of eosinophils, macrophages, neutrophils, and Tem compared to Cluster 2. The xCell algorithm was utilized to quantify immune cell infiltration, as demonstrated in Figure 9B. Cluster 1 consistently exhibited heightened levels of B cells, CD8+ T cells, Th1 cells, and Th2 cells, indicating a notable disparity compared to Cluster 2. Conversely, Cluster 1 displayed reduced infiltration of neutrophils and NKT cells. Figure 9C demonstrates the ssGSEA scores for various immune-related pathways, with notable emphasis on antimicrobials, chemokine receptors, chemokines, cytokine receptors, cytokines, interleukins, natural killer cell cytotoxicity, and TNF family members. Despite the lower level of cell infiltration in Cluster 2, these pathways exhibit heightened activity compared to Cluster 1. This increased activity in Cluster 2 may indicate a more robust immune signaling pathway. In summary, the comparison of BLI-related ASD subgroups reveals a paradox in which Cluster 2 exhibits heightened immune pathway activities alongside diminished immune cell infiltration. This discovery suggests a complex regulatory milieu in Cluster 2, potentially indicating a unique pathophysiology within the immune landscape of this subgroup.

Correlation analysis of bacterial lipoproteins pathway activity with immune cell infiltration and immune-related pathways in ASD subgroups

In analyzing subgroups of ASD associated with BLI, we investigated the activity of the bacterial lipoproteins pathway, along with its correlation with immune cell infiltration and immune-related pathways. Figure 10A and C present GSVA scores for three different immune response pathways: response to bacterial lipoprotein, cellular response to bacterial lipoprotein, response to diacyl bacterial lipoprotein. The violin plots reveal that Cluster 2 exhibits significantly higher activity in these pathways when compared to Cluster 1, implicating a more pronounced immunological response to bacterial components within this subgroup. Figure 10D is a heat map which represents the correlation coefficients between the GSVA scores of the bacterial lipoproteins pathways and the immune cell infiltrate scores for various immune cells. Notably, several immune cell types, such as NK cells, M1 macrophages, monocytes and neutrophils, show strong positive correlations with these pathway activities. Similarly, Figure 10E depicts the heat map of correlation coefficients between the bacterial lipoproteins pathways and various immune-related pathways. A strong positive correlation is seen with pathways like cytokine signaling and the interferon response. This highlights the potential interaction of these pathways with the heightened immune activity against bacterial components, suggesting a complex interplay in the immunological landscape of this ASD subgroup. Overall, the data emphasizes a heightened state of immune-related pathway activity and a specific but lower level of immune cell infiltration in the BLI-related Cluster 2. The correlations suggest a dynamic association between pathway activities and the composition of the immune cell infiltrate, which may be crucial in understanding the underlying immunological mechanisms in ASD.

Evaluation of Key BLI Gene Expression

To validate the expression of key genes associated with BLI in ASD, clinical samples were collected and analyzed. As shown in Figure 11, the BLNK shows a substantial downregulation in ASD samples compared to HC. In contrast, the MAPK8, PRKCQ, and TNFSF12 exhibit a pronounced upregulation in ASD individuals relative to HC.

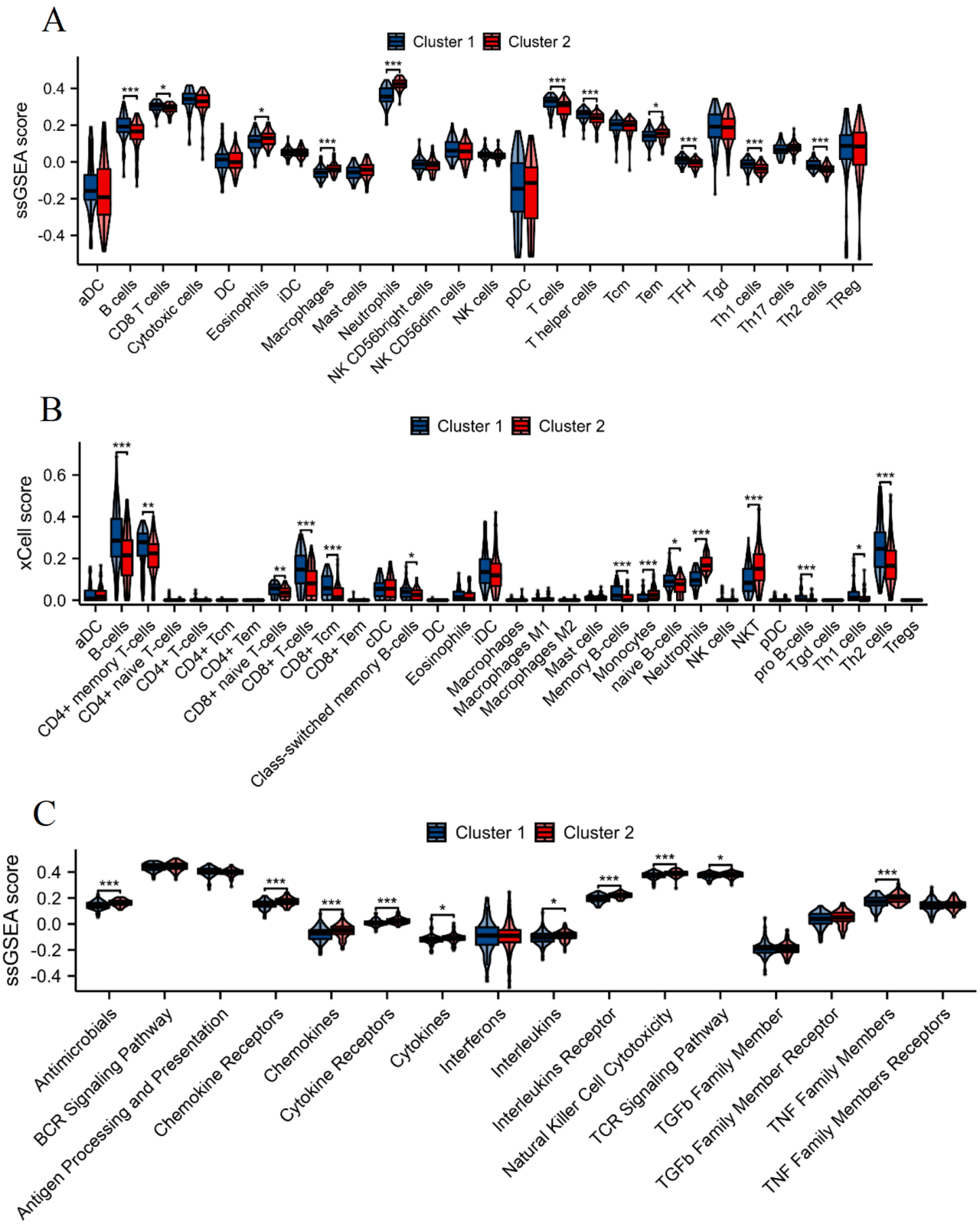

Figure 9 Comparative analysis of immune cell infiltration and pathway activity in BLI-related ASD subgroups. The violin plots represent the scores derived from ssGSEA (**A**) and xCell (**B**) for a range of immune cells. (**C**) The violin plots depict ssGSEA-derived scores assessing the activity of different immune-related pathways. *p < 0.05, **p < 0.01, ***p < 0.001.

https://doi.org/10.2147/PGPM.S494126

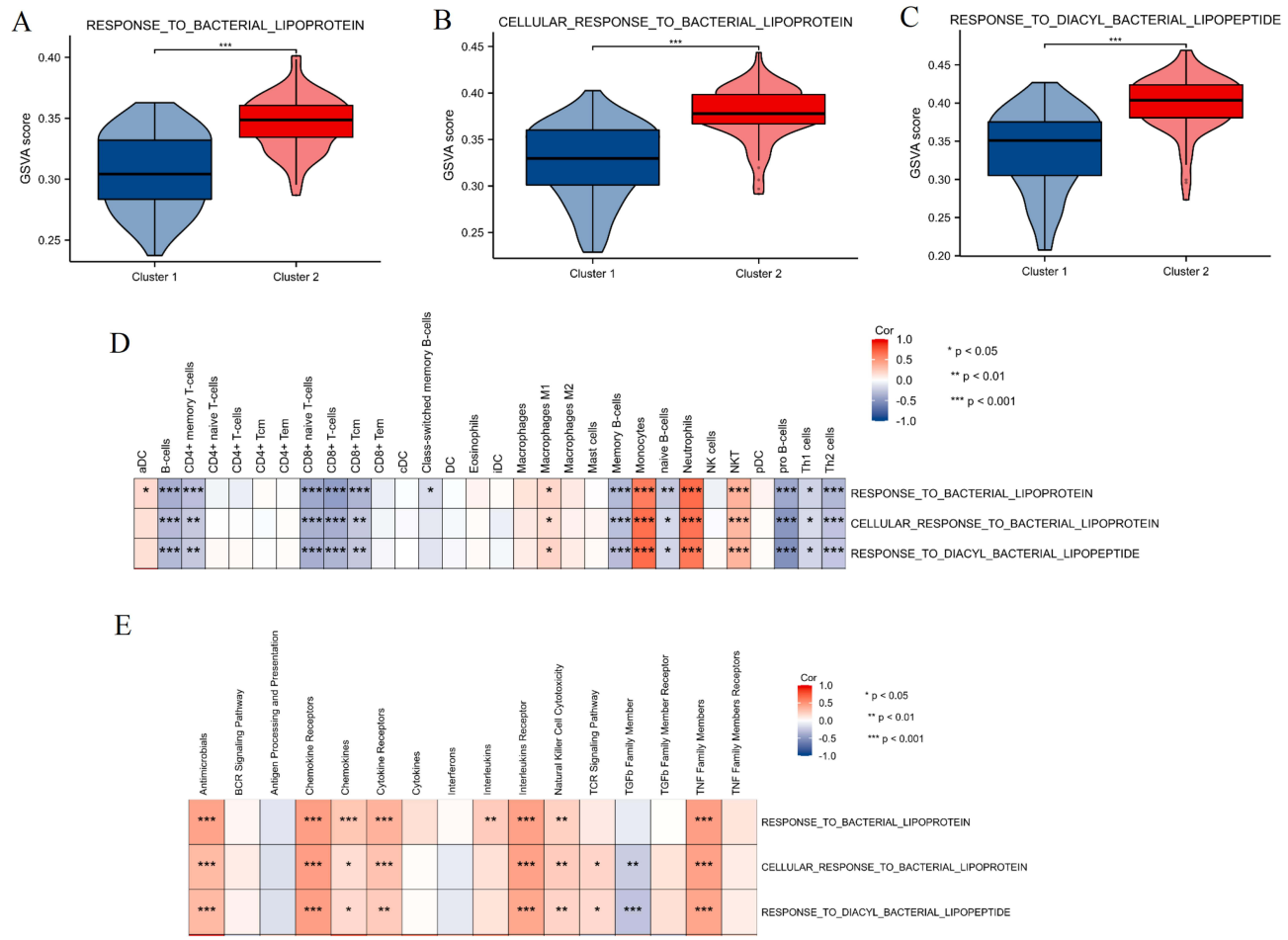

Figure 10 Association between bacterial lipoprotein pathway activity and immune cell infiltration. This violin plot illustrates GSVA scores comparing the activity of response to bacterial lipoprotein (**A**), cellular response to bacterial lipoprotein (**B**), response to diacyl bacterial lipoprotein (**C**) in Cluster 1 (blue) versus Cluster 2 (red) of ASD subgroups. (**D**) A heat map visualizes the correlation between the activities of bacterial lipoproteins pathways and the infiltration scores of various immune cell types. (**E**) The heat map depicts the correlation coefficients between bacterial lipoproteins pathway activities and different immune-related pathways. *p < 0.05, **p < 0.01, ***p < 0.001.

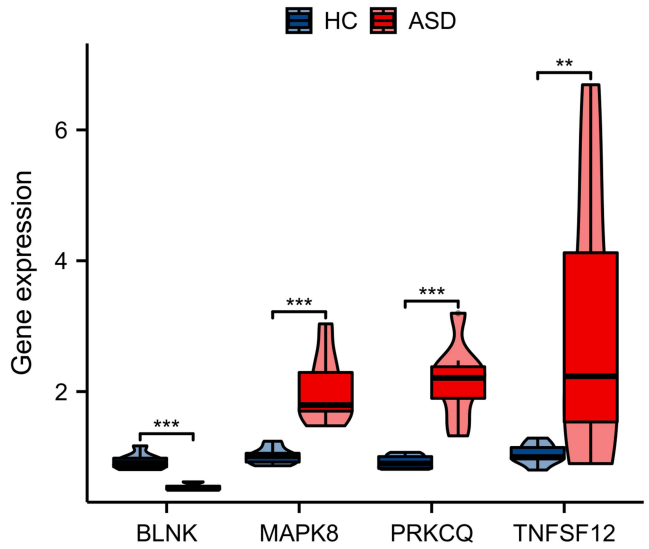

Figure 11 Validation of key BLI gene expression. **p < 0.01, ***p < 0.001.

Discussion

The intricate relationship between the immune system and neurodevelopmental disorders such as ASD is increasingly recognized as a multifaceted contributor to the etiology and clinical diversity of ASD.[19] Our study has moved the needle forward in understanding this complex relationship by elucidating the significance of BLI genes in the classification of ASD. Prior literature has documented the potential role of immune dysregulation and microbial components in ASD,[20–22] and our findings corroborate and expand upon these previous insights.

The identification of 17 BLI genes exhibiting differential expression patterns in ASD, notably BLNK, MAPK8, PRKCQ, and TNFSF12, resonates with previous immunological studies in human diseases. For instance, BLNK has been implicated in B cell development and function,[23,24] and the dysregulation we observed suggests a possible aberration in humoral immunity in ASD. Similarly, the upregulation of MAPK8, a gene associated with the MAPK signaling pathway, provides support that dysregulation of this critical pathway may contribute to the pathogenesis of autism.[25,26] The establishment of a nomogram prediction model for ASD based on BLI signatures presents both novelty and practical implications for clinical assessment and diagnosis. Our model's high diagnostic accuracy highlights its potential utility as a non-invasive, predictive tool, capable of aiding early diagnosis and possibly tracking the effectiveness of interventions.

The elucidation of molecular subtypes within ASD holds significant promise for understanding its etiology and advancing personalized treatments. The identification of two distinct clusters suggests ASD is not a single entity but rather a spectrum of disorders with heterogeneous biological basis. Cluster 1 may represent a subtype less influenced by immune dysfunction, whereas Cluster 2, characterized by upregulation of immune and inflammation-related pathways, could suggest a more biologically distinct group possibly amenable to targeted interventions. The heightened expression of immune markers in Cluster 2 subgroup, including cytokines and chemokines, along with increased natural killer cell cytotoxicity, indicates an active neuroinflammatory process. Neuroinflammation has been proposed as a contributing factor to the development of ASD, with several studies noting altered immune responses in these individuals.[27–29] The heightened state of immune activation in Cluster 2 could be a response to various environmental insults, including prenatal infections or postnatal immune challenges, which have been associated with an increased risk of ASD.[30,31] It is also worth considering that gender differences may contribute to variability in immune responses and clustering patterns, as ASD has been shown to manifest differently in males and females, possibly influencing immune activation. The elevated immune markers may also reflect an intrinsic dysregulation of the immune system, potentially stemming from genetic susceptibilities that affect immune function and neurodevelopment.[32] The presence of elevated antimicrobial agents might also be indicative of an ongoing attempt by the immune system to counteract microbial or viral infections, which some hypotheses suggest could play a role in the development of ASD.[33] However, it is also possible that this is a secondary effect, with the immune system activation being a consequence of a primary neurodevelopmental process that is disturbed in ASD. Therapeutically, these insights hold substantial implications. If the neuroinflammation observed in Cluster 2 is a driving force in the pathophysiology, treatments aimed at modulating the immune system might have a beneficial effect on the symptoms of ASD in this subtype. For instance, the use of anti-inflammatory drugs, immunomodulatory interventions, or even dietary changes aimed at reducing inflammation could be potential strategies for amelioration of symptoms.[34,35]

The elevated levels of immune cell infiltration in Cluster 1, encompassing various lymphocyte subsets such as B cells, T cells, and specific T cell subtypes, suggest a heightened adaptive immune response. This may imply a dysregulated immune activation in Cluster 1, potentially driven by immune dysregulation, which has been implicated in the pathogenesis of ASD.[8,29] Conversely, the reduced infiltration of certain immune cell populations, such as neutrophils and NKT cells, in Cluster 1 indicates a distinct immunological profile characterized by altered cellular immune responses. The contrasting immune landscape in Cluster 2, characterized by heightened immune pathway activities despite diminished immune cell infiltration, presents a paradoxical situation. This discrepancy may imply a dysregulated immune signaling cascade in Cluster 2, where aberrant cytokine or chemokine signaling pathways drive immune activation independently of immune cell abundance. Such dysregulated immune signaling pathways have been implicated in neuroinflammatory processes and synaptic dysfunction associated with ASD.[36]

In addition, the activation of bacterial lipoprotein-related pathways and its association with immune cell presence represent a fascinating avenue for exploring the potential role of microbial triggers in ASD pathogenesis. The link between bacterial

lipoproteins and immune responses suggests a possible mechanism whereby dysbiosis or unresolved bacterial infections could perturb the gut microbiota and subsequently elicit aberrant immune activation, with implications for neurological function and behavior in individuals with ASD. The gut-brain axis, a bidirectional communication network between the gastrointestinal tract and the central nervous system, has garnered increasing attention in the context of neurodevelopmental disorders such as ASD.[37,38] Dysbiosis, characterized by alterations in the composition and function of the gut microbiota, has been observed in a subset of individuals with ASD.[39] Perturbations in the gut microbiota can lead to the dysregulation of immune responses, as the gut is a major site of interaction between commensal microbes and the host immune system.[40] Bacterial lipoproteins, which are components of the cell wall in many Gram-positive and Gram-negative bacteria, have been implicated as potent activators of the innate immune system.[41] Recognition of bacterial lipoproteins by pattern recognition receptors, such as Toll-like receptors, triggers signaling cascades that culminate in the production of pro-inflammatory cytokines and chemokines, as well as the recruitment of immune cells to the site of infection.[42] The activation of bacterial lipoprotein-related pathways in ASD may therefore reflect a dysregulated immune response to microbial stimuli. Furthermore, the association between bacterial lipoprotein-related pathways and immune cell presence suggests a dynamic interplay between microbial triggers, immune activation, and neuroinflammation in ASD pathophysiology. Chronic immune activation and neuroinflammation have been implicated in the pathogenesis of ASD, with evidence of increased levels of pro-inflammatory cytokines and activated microglia in the brains of individuals with ASD.[29,43] The activation of immune responses by microbial triggers could exacerbate neuroinflammation and contribute to synaptic dysfunction, altered neural connectivity, and behavioral abnormalities observed in ASD. Additionally, it would be valuable to consider the severity of infection or immune challenges in different patient populations, as this may influence the immune response and clustering patterns observed in our study.

It is important to note that while the activation of bacterial lipoprotein-related pathways in ASD may implicate a role for microbial triggers in disease pathogenesis, further research is needed to elucidate the specific microbial species or communities involved, as well as the mechanisms underlying immune dysregulation and neuroinflammation. In summary, the activation of bacterial lipoprotein-related pathways and its association with immune cell presence offer clues to the potential role of microbial triggers in immune dysregulation and neuroinflammation in ASD. Understanding the complex interactions between the gut microbiota, the immune system, and the central nervous system may ultimately lead to novel therapeutic interventions targeting the gut-brain axis in ASD.

Conclusion

In conclusion, our study highlights the significance of genes associated with LPS and immune response in the classification of ASD. Notably, some of these genes are also involved in glial regulation, underscoring their potential role in neurodevelopmental processes. Furthermore, our findings revealed two distinct molecular subtypes, with Cluster 2 displaying upregulation of immune- and inflammation-related pathways. Overall, our research underscores the intricate interplay between bacterial components, genetic predisposition, and immune dysregulation in ASD. These findings not only contribute to a better understanding of the complexities of ASD but also offer insights into the development of personalized interventions for individuals with this neurodevelopmental disorder.

Data Sharing Statement

All data used in the present study were available from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). The accession numbers are as follows: GSE18123 and GSE42133).

Ethics Approval

This research was carried out in alignment with the Declaration of Helsinki and received approval from the Ethics Committee at North Sichuan Medical College. Additionally, informed consent was secured from all participants involved.

Consent for Publication

Not applicable.

Acknowledgment

This work was supported by the North Sichuan Medical College Innovation Center for Science and Technology.

Funding

This work was supported by Natural Science Foundation of Sichuan Province (2024NSFSC1670), Key Cultivation Program of North Sichuan Medical College (CBY23-ZDA05) and the PhD Research Fund of Affiliated Hospital of North Sichuan Medical College (202004).

Disclosure

All authors declared that there was no conflicts of interest in this work.

References

1. Hirota T, King BH. Autism spectrum disorder: a review. *JAMA*. 2023;329(2):157–168. doi:10.1001/jama.2022.23661
2. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. *Lancet (London, England)*. 2018;392(10146):508–520. doi:10.1016/S0140-6736(18)31129-2
3. Masi A, DeMayo MM, Glozier N, Guastella AJ. An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience Bulletin*. 2017;33(2):183–193. doi:10.1007/s12264-017-0100-y
4. Wang L, Wang B, Wu C, Wang J, Sun M. Autism spectrum disorder: neurodevelopmental risk factors, biological mechanism, and precision therapy. *Int J Mol Sci*. 2023;24(3):1819.
5. Baumer N, Spence SJ. Evaluation and management of the child with autism spectrum disorder. *Continuum*. 2018;24(1, Child Neurology):248–275. [(1, Child Neurology)]. doi:10.1212/CON.0000000000000578
6. Al-Jabri Z, Al-Reesi I, Al-Shizawi N, Al-Balushi MS, Al-Jabri AA, Said EA. Effects of bacterial metabolites on the immune system: enemies and friends. *Endocr. Metab. Immune Disord. Drug Targets*. 2022;22(12):1167–1177. doi:10.2174/1871530321666211119150231
7. Gasaly N, de Vos P, Hermoso MA. Impact of bacterial metabolites on gut barrier function and host immunity: a focus on bacterial metabolism and its relevance for intestinal inflammation. *Front Immunol*. 2021;12:658354. doi:10.3389/fimmu.2021.658354
8. Estes ML, McAllister AK. Immune mediators in the brain and peripheral tissues in autism spectrum disorder. *Nat Rev Neurosci*. 2015;16 (8):469–486. doi:10.1038/nrn3978
9. Theoharides TC, Tsilioni I, Patel AB, Doyle R. Atopic diseases and inflammation of the brain in the pathogenesis of autism spectrum disorders. *Transl Psychiatry*. 2016;6(6):e844. doi:10.1038/tp.2016.77
10. Hsiao EY, McBride SW, Hsien S, et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell*. 2013;155(7):1451–1463. doi:10.1016/j.cell.2013.11.024
11. Eicher TP, Mohajeri MH. Overlapping mechanisms of action of brain-active bacteria and bacterial metabolites in the pathogenesis of common brain diseases. *Nutrients*. 2022;14(13):2661. doi:10.3390/nu14132661
12. Careaga M, Rogers S, Hansen RL, Amaral DG, Van de Water J, Ashwood P. Immune endophenotypes in children with autism spectrum disorder. *Biol. Psychiatry*. 2017;81(5):434–441. doi:10.1016/j.biopsych.2015.08.036
13. Suprunowicz M, Tomaszek N, Urbaniak A, Zackiewicz K, Modzelewski S, Waszkiewicz N. Between Dysbiosis, Maternal Immune Activation and Autism: is There a Common Pathway? *Nutrients*. 2024;16(4):549. doi:10.3390/nu16040549
14. Kaminski VL, Michita RT, Ellwanger JH, et al. Exploring potential impacts of pregnancy-related maternal immune activation and extracellular vesicles on immune alterations observed in autism spectrum disorder. *Heliyon*. 2023;9(5):e15593. doi:10.1016/j.heliyon.2023.e15593
15. Li S, Ma L, Cui R. Identification of novel diagnostic biomarkers and classification patterns for osteoarthritis by analyzing a specific set of genes related to inflammation. *Inflammation*. 2023;46(6):2193–2208. doi:10.1007/s10753-023-01871-w
16. Liang A, Feng T, Gao X, Zhao B, Chen S. Identification of PDGFA as a neutrophil-related biomarker linked to the advancement of diabetic retinopathy through integrated bioinformatics analysis. *Endocr. Metab. Immune Disord. Drug Targets*. 2024.
17. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102(43):15545–15550. doi:10.1073/pnas.0506580102
18. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;14(1):7. doi:10.1186/1471-2105-14-7
19. Hughes HK, Moreno RJ, Ashwood P. Innate immune dysfunction and neuroinflammation in autism spectrum disorder (ASD). *Brain Behav Immun*. 2023;108:245–254. doi:10.1016/j.bbi.2022.12.001
20. Srikantha P, Mohajeri MH. The possible role of the microbiota-gut-brain-axis in autism spectrum disorder. *Int J Mol Sci*. 2019;20(9):2115. doi:10.3390/ijms20092115
21. Jyonouchi H, Sun S, Le H. Proinflammatory and regulatory cytokine production associated with innate and adaptive immune responses in children with autism spectrum disorders and developmental regression. *J Neuroimmunol*. 2001;120(1–2):170–179. doi:10.1016/S0165-5728(01)00421-0
22. Robinson-Agramonte MLA, Noris García E, Fraga Guerra J, et al. Immune dysregulation in autism spectrum disorder: what do we know about it? *Int J Mol Sci*. 2022;23(6):3033. doi:10.3390/ijms23063033
23. Tsukada S, Baba Y, Watanabe D. Btk and BLNK in B cell development. *Adv Immunol*. 2001;77:123–162.
24. Xu S, Tan JE, Wong EP, Manickam A, Ponniah S, Lam KP. B cell development and activation defects resulting in xid-like immunodeficiency in BLNK/SLP-65-deficient mice. *Int Immunol*. 2000;12(3):397–404. doi:10.1093/intimm/12.3.397
25. Vithayathil J, Pucilowska J, Landreth GE. ERK/MAPK signaling and autism spectrum disorders. *Prog Brain Res*. 2018;241:63–112. doi:10.1016/bs.pbr.2018.09.008

26. Garg S, Brooks A, Burns A, et al. Autism spectrum disorder and other neurobehavioural comorbidities in rare disorders of the Ras/MAPK pathway. *Dev Med Child Neurol*. 2017;59(5):544–549. doi:10.1111/dmcn.13394
27. Xiong Y, Chen J, Li Y. Microglia and astrocytes underlie neuroinflammation and synaptic susceptibility in autism spectrum disorder. *Front Neurosci*. 2023;17:1125428. doi:10.3389/fnins.2023.1125428
28. Matta SM, Hill-Yardin EL, Crack PJ. The influence of neuroinflammation in autism spectrum disorder. *Brain Behav Immun*. 2019;79:75–90. doi:10.1016/j.bbi.2019.04.037
29. Ashwood P, Krakowiak P, Hertz-Picciotto I, Hansen R, Pessah IN, Van de Water J. Altered T cell responses in children with autism. *Brain Behav Immun*. 2011;25(5):840–849. doi:10.1016/j.bbi.2010.09.002
30. Atladóttir HO, Thorsen P, Østergaard L, et al. Maternal infection requiring hospitalization during pregnancy and autism spectrum disorders. *J Autism Dev Disord*. 2010;40(12):1423–1430. doi:10.1007/s10803-010-1006-y
31. Jiang HY, Xu LL, Shao L, et al. Maternal infection during pregnancy and risk of autism spectrum disorders: a systematic review and meta-analysis. *Brain Behav Immun*. 2016;58:165–172. doi:10.1016/j.bbi.2016.06.005
32. Børglum AD, Demontis D, Grove J, et al. Genome-wide study of association and interaction with maternal cytomegalovirus infection suggests new schizophrenia loci. *Mol Psychiatry*. 2014;19(3):325–333. doi:10.1038/mp.2013.2
33. Hornig M, Bresnahan MA, Che X, et al. Prenatal fever and autism risk. *Mol Psychiatry*. 2018;23(3):759–766. doi:10.1038/mp.2017.119
34. Theoharides TC, Asadi S, Patel AB. Focal brain inflammation and autism. *J Neuroinflammation*. 2013;10(1):46. doi:10.1186/1742-2094-10-46
35. Gesundheit B, Rosenzweig JP, Naor D, et al. Immunological and autoimmune considerations of Autism Spectrum Disorders. *J Autoimmun*. 2013;44:1–7. doi:10.1016/j.jaut.2013.05.005
36. Di Marco B, Bonaccorso CM, Aloisi E, D'Antoni S, Catania MV. Neuro-inflammatory mechanisms in developmental disorders associated with intellectual disability and autism spectrum disorder: a neuro- immune perspective. *CNS & Neurological Disord Drug Targets*. 2016;15(4):448–463. doi:10.2174/1871527315666160321105039
37. Mayer EA, Knight R, Mazmanian SK, Cryan JF, Tillisch K. Gut microbes and the brain: paradigm shift in neuroscience. *J Neurosci*. 2014;34(46):15490–15496. doi:10.1523/JNEUROSCI.3299-14.2014
38. Dargenio VN, Dargenio C, Castellaneta S, et al. Intestinal barrier dysfunction and microbiota-gut-brain axis: possible implications in the pathogenesis and treatment of autism spectrum disorder. *Nutrients*. 2023;15(7):1620. doi:10.3390/nu15071620
39. Kang DW, Adams JB, Gregory AC, et al. Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome*. 2017;5(1):10. doi:10.1186/s40168-016-0225-7
40. Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell*. 2014;157(1):121–141. doi:10.1016/j.cell.2014.03.011
41. Kovacs-Simon A, Titball RW, Michell SL. Lipoproteins of bacterial pathogens. *Infect Immun*. 2011;79(2):548–561. doi:10.1128/IAI.00682-10
42. Kumar V. Toll-like receptors in adaptive immunity. *Handbook Exp Pharmacol*. 2022;276:95–131.
43. Vargas DL, Nascimbene C, Krishnan C, Zimmerman AW, Pardo CA. Neuroglial activation and neuroinflammation in the brain of patients with autism. *Ann. Neurol.* 2005;57(1):67–81. doi:10.1002/ana.20315