Immune-related Signature of Periodontitis and Alzheimer’s disease linkage

Jieqi Jin
Beijing Friendship Hospital

Mengkai Guang
China-Japan Friendship Hospital

Simin Li
Southern Medical University

Xing Yan
Beijing Friendship Hospital

Yong Liu
Beijing Friendship Hospital

Liwei Zhang
Beijing Friendship Hospital

Bo Zhang
Beijing Friendship Hospital

Menglin Cheng
Beijing Friendship Hospital

Gerhard Schmalz
Leipzig University

Xiaofeng Huang (huangxf1998@163.com)
Beijing Friendship Hospital

Research Article

Keywords: Alzheimer’s disease, Periodontitis, Immune cell, Gene-signature

Posted Date: May 15th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2907935/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

Additional Declarations: No competing interests reported.
Immune-related Signature of Periodontitis and Alzheimer’s disease linkage

Jieqi Jin1†, Mengkai Guang2†, Simin Li3, Xing Yan1, Yong Liu1, Liwei Zhang1, Bo Zhang1, Menglin Cheng1, Gerhard Schmalz4#, Xiaofeng Huang1* 

1Department of Stomatology, Beijing Friendship Hospital, Capital Medical University, Yongan Road No.95, Beijing 100050, China.
2Department of Stomatology, China-Japan Friendship Hospital, Yinghua Street No.2, Beijing 100029, China.
3Stomatological Hospital, Southern Medical University, Jiangnan Road South No.366, Guangzhou 510280, China.
4Department of Cariology, Endodontology and Periodontology, Leipzig University, Liebigstr. 10-14, 04103, Leipzig, Germany 

*Corresponding author. E-mail: huangxf1998@163.com;
#Senior author. E-mail: Gerhard.Schmalz@medizin.uni-leipzig.de
†These authors contributed equally to this work.

Abstract

Background: Periodontitis (PD) and Alzheimer’s disease (AD) are both associated with ageing and clinical studies increasingly evidence their association. However, specific mechanisms underlying this association remain undeciphered, and immune-related processes are purported to play a significant role. The accrual of publically available transcriptomic datasets permits secondary analysis and the application of data-mining and bioinformatic tools for biological discovery. Aim: The present study aimed to leverage publically available transcriptomic datasets and databases, and apply a series of bioinformatic analysis to identify a robust signature of immune-related signature of PD and AD linkage. Methods: We downloaded gene-expression data pertaining PD and AD and identified crosstalk genes. We constructed a protein-protein network analysis, applied immune cell enrichment analysis, and predicted crosstalk immune-related genes and infiltrating immune cells. Next, we applied consisent
cluster analysis and performed immune cell bias analysis, followed by LASSO regression to select biomarker immune-related genes. **Results:** The results showed a 3 gene set comprising of DUSP14, F13A1 and SELE as a robust immune-related signature. Macrophages M2 and NKT, B-cells, CD4+ memory T-cells and CD8+ naive T-cells emerged as key immune cells linking PD with AD. **Conclusion:** Candidate immune-related biomarker genes and immune cells central to the association of PD with AD were identified, and merit investigation in experimental and clinical research.

**Keywords:** Alzheimer’s disease, Periodontitis, Immune cell, Gene-signature.

## 1 Introduction

With the rapid ageing of global populations, the burden of Alzheimer’s disease (AD) is rising. AD is a neurodegenerative disease marked by the formation of amyloid-β peptide (AβP) plaques aggregate in brain tissues. Inflammation and pathological aberrations in central and peripheral immune responses are implicated in AD[1, 2]. While the relationship of systemic or peripheral inflammation with AD has been inconsistent [3], accruing research has highlighted the role of peripheral inflammation in AD pathogenesis. Gut microbiome dysbiosis is associated with neuroinflammation and synaptic dysfunction characteristic of AD [4]. Research using a murine model of AD has demonstrated that low-grade peripheral inflammation is capable of aggravating brain pathology[5]. ApoE4 allele of the Apolipoprotein E gene, a well-known genetic risk factor of AD, when coupled with chronic low-grade peripheral inflammation leads to earlier onset and greater morbidity from AD[6]. Peripheral inflammation also alters the connectivity of large-scale cognitive networks in older individuals, particularly in ApoE4 carriers[7]. Elevated levels of both peripheral and CSF inflammatory markers are associated with AD[8]. Systemic infections are associated with enhanced immunosuppressive processes in the brains of patients with AD, with increase in anti-inflammatory proteins including IL4R and CHI3L1 and decrease in certain proinflammatory proteins, along with lowered T-cell recruitment [9]. Systemic inflammation can affect intra-brain drug distribution by altering ABCB1 and ABCG2 protein expression and perturb GluN1 protein expression in AD affected brains [10]. Circulating IL-21, a key immunomodulatory cytokine is elevated in AD, possibly due to immune activation and results in neuroinflammation, microglial activation, and deposition of Aβ plaques[11].

Ageing is associated with higher levels of chronic inflammation and immune deregulation.
Infections cause immune dysregulation, increase circulating pro-inflammatory mediators such as TNFα and IL-6 along with the brain levels of IL-1β and IL-6 levels, aggravating neuroinflammation and accelerating cognitive decline in older adults [12, 13]. Immune perturbation in AD is not restricted to the central nervous system (CNS), and peripheral immune dysregulation appears to affect homeostasis in AD affected brains, where the barrier function is disrupted, allowing an ingress of T-cells [14]. Perturbed naive and memory CD4+ T cell subsets have been noted in the peripheral blood of patients with mild AD and dementia, with a lower proportion of naive cells and an increased proportion of effector memory and terminal differentiation effector memory (TEMRA) CD4+ cells[15]. A deregulation of both peripheral and central immune compartments marks AD. Peripheral immune activation is associated with neuroinflammation and AD pathogenesis. Sustained activation of the brain’s microglia and other immune cells is found to exacerbate both amyloid and tau pathology, and may serve as the link between infections, chronic peripheral inflammation and AD [16].

Periodontitis is a highly prevalent oral infectious disease that imposes both oral and systemic health burden. It is an inflammatory disease caused by a complex interplay between dental plaque microbes and the host immune system[17]. The deposition of a microbial plaque biofilm initiates immune cell migration, and its dysbiosis sustains a local inflammatory response[18]. Key periodontal pathogens such as Porphyromonas gingivalis are immune evasive and can activate the complement system and pathogen recognition receptors such as TLRs, leading to chronic inflammation and periodontal tissue destruction[19]. Virulence factors like Porphyromonas gingivalis fimbriae can activate TLR2 expressed by innate immune cells[20], leading to a cascade of cellular and humoral immune responses, and induction of adaptive immune responses [17]. Ageing is associated with a steep increase in the incidence and severity of periodontitis, attributed in part to increased susceptibility from age-dependent alterations in host innate immunity and inflammatory status[21]. Cellular senescence, stem cell failure, and immune senescence inherent to biological ageing impair periodontal tissue homeostasis and contribute to the pathophysiology of periodontitis[22].

Evidence showing the association of AD with periodontitis is rapidly accumulating [23, 24, 25]. Porphyromonas gingivalis has been found to infiltrate the brains of tissue in AD and is proposed to be an important mechanistic link between periodontitis and AD[26]. Periodontitis also causes widespread
systemic immune dysfunction, showing heightened pro-inflammatory responses to *Porphyromonas gingivalis* and attenuated T-cell responses [27]. In the present study, we aimed to identify immunological perturbations and immune crosstalk potentially linking periodontitis with AD by leveraging gene expression data.

2 Materials and methods
2.1 Datasets
We downloaded expression profile datasets related to periodontitis (PD) and Alzheimer’s disease (AD) from the GEO ([https://www.ncbi.nlm.nih.gov/geo/](https://www.ncbi.nlm.nih.gov/geo/)) database. For PD, we chose gingival tissue; for AD, we chose brain tissue data. The datasets are listed in table 1.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Datasets</th>
<th>Platform</th>
<th>Case</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>GSE10334</td>
<td>GPL570</td>
<td>241</td>
<td>69</td>
<td>310</td>
</tr>
<tr>
<td></td>
<td>GSE16134</td>
<td>GPL570</td>
<td>183</td>
<td>64</td>
<td>247</td>
</tr>
<tr>
<td></td>
<td>GSE106090</td>
<td>GPL21827</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>AD</td>
<td>GSE33000</td>
<td>GPL4372</td>
<td>310</td>
<td>157</td>
<td>467</td>
</tr>
<tr>
<td></td>
<td>GSE36980</td>
<td>GPL6244</td>
<td>33</td>
<td>47</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>GSE122063</td>
<td>GPL16699</td>
<td>56</td>
<td>44</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>GSE48350</td>
<td>GPL570</td>
<td>80</td>
<td>173</td>
<td>253</td>
</tr>
<tr>
<td></td>
<td>GSE5281</td>
<td>GPL570</td>
<td>87</td>
<td>74</td>
<td>161</td>
</tr>
</tbody>
</table>
2.2 Differential gene expression analysis

First, we converted the probe names into gene names based on the downloaded data. If the same gene had multiple expression values in the same sample, we obtained the mean of the expression values. As differences existed between the datasets, we first combined the datasets for AD and PD each based on common genes, and then applied the ‘ComBat’ method in the R package ‘SVA’ for batch correction. Among the datasets related to AD, since the series matrix of GSE33000 was a standardised dataset, the other four datasets were standardised separately. We then combined the 5 standardised datasets and applied the ‘ComBat’ method to perform batch correction.

Differential expression gene analysis of the corrected datasets was performed using the R package ‘limma’. For the AD datasets, we used a threshold of p value < 0.05, with |log2 (FC)| > 0 for upregulated genes and log2 (FC) < 0 for downregulated genes. For PD datasets we used p value < 0.05, log2 (FC) > 0.5 for upregulated genes and log2 (FC) ≤ 0.5 for downregulated genes.

2.3 Identification of crosstalk genes

The differentially expressed genes of AD and PD were intersected and the shared genes were regarded as potential crosstalk genes. Functional enrichment analysis of the crosstalk genes was performed using ‘clusterProfiler’ (GO Biological processes and KEGG pathways, at a threshold of p value<0.05.

2.4 Crosstalk genes’ protein-protein interaction (PPI) network

We downloaded protein-protein related gene pairs from MINT (http://mint.bio.uniroma2.it/mint/Welcome.do), HPRD (http://www.hprd.org/index_html), BIOGRID (http://thebiogrid.org/), DIP (http://dip.doe-mbi.ucla.edu/dip/Main.cgi), mentha (http://mentha.uniroma2.it/index.php), PINA (http://cbg.garvan.unsw.edu.au/pina/), InnateDB (http://www.innatedb.com/), and Instruct (http://instruct.yulab.org/index.html) databases. Next, PPI relationship pairs for the crosstalk gene were extracted, and a PPI network was constructed using Cytoscape software, with the plug-in NetworkAnalyzer to analyse the network topological properties.

2.5 Crosstalk genes’ immune cell enrichment analysis
XCell (https://xcell.ucsf.edu/) was used for cell type enrichment of the crosstalk genes. XCell includes 64 cell types involving multiple adaptive and innate immune cells, hematopoietic progenitors, epithelial cells, and extracellular matrix cells, which includes 48 tumor microenvironment-related cells. We first extracted Case samples from AD and PD datasets and obtained the expression values of crosstalk genes. The gene number limit for the raw enrichment analysis method in the xCell package was reset and the scores for immune-infiltrating cells corresponding to the samples were calculated. Next, ‘transform scores’ and ‘spill over’ were used to obtain the final corrected immune infiltrating cell scores.

2.6 Analysis of immune-related genes and infiltrating immune cells

Immune-related genes were downloaded from an earlier publication [28] (782 genes, including 431 genes related to 15 adaptive immune cell types and 351 genes related to 13 innate immune cell types) and combined with immune-related genes from Innate DB (https://innatedb.com/annotatedGenes.do?type=innatedb) and ImmPort (https://www.immport.org/home) datasets. The expression values of these immune-related genes in the AD and PD case samples were extracted and the xCell algorithm was used to obtain the expression scores of infiltrating immune cells. The relationship between the immune cells in the two diseases was tested using the Wilcoxon’s test (P <0.05).

2.7 Consistent cluster analysis of AD and PD samples based on immune-related genes

We applied consensus clustering to the expression matrix profiles of immune-related genes in AD and PD each, using the ‘ConsensusClusterPlus’ package. Average silhouette width, gap statistic, and the elbow method were used to determine the optimal number of clusters. Next, the samples were clustered using clustering consistency.

2.8 Adaptive and innate immune cell bias analysis

We combined the immune cell fractions and sample clusters of case samples from AD and PD. For each consensus cluster, statistics for cluster distribution of xCell scores for adaptive and innate immune-related genes were computed. The difference between scores of immune cells in different clusters was tested using the Kruskal Wallis test. We also noted the overall scores of immune cells in different clusters to determine immune cells that characterised a cluster.
2.9 Identification of Potential Biomarkers Using LASSO Logistic Regression

We extracted the expression values of the crosstalk genes for the case and control groups and applied LASSO logistic regression. From the screened crosstalk genes, those common to AD and PD were considered biomarker crosstalk genes. Next, adaptive immune cell-related genes were identified based on literature and their expression profiles in AD and PD datasets were screened using LASSO logistic regression. The intersecting genes were recorded as biomarker adaptive immune genes. For Innate immune cell-related genes, we combined those obtained from literature with those obtained from the InnateDB dataset and obtained 1335 innate immune-cell related genes. LASSO logistic regression was similarly applied and biomarker innate immune genes were identified. In the next step, the intersections of biomarker crosstalk genes with biomarker adaptive immune cell-related genes and biomarker innate immune cell-related genes were determined. Receiver operating curve (ROC) analysis was performed using these genes’ expression values. Human KEGG pathways and related genes were obtained from the KEGG database (https://www.kegg.jp/) and pathways that correspond to these intersecting genes were identified and all genes in each such pathway were listed. Interactions between the KEGG pathways, biomarker crosstalk genes, biomarker adaptive immune-cell related genes, and biomarker innate immune cell-related genes were identified.

3 Results

3.1 Differentially expressed genes

Combining the AD and PD related datasets each and applying batch correction, differences noted before correction were significantly reduced post batch correction (Fig.1) (file1 file2).
Differential expression analysis of using the batch corrected data (file3-6) we obtained 4398 differentially expressed genes in AD and 1041 differentially expressed genes in PD. A volcano map was used to display the distribution of the differentially expressed genes (Fig 2).
3.2 Crosstalk genes

A total of 364 Crosstalk genes were obtained (Fig 3A) (file7). To visualize changes in the expression values of crosstalk genes in different sample types, heat maps were plotted using the ‘pheatmap’ R package and 50 Crosstalk genes were displayed (file8-9)(Fig 3B-C). To further analyze the functions of the crosstalk genes functional enrichment analysis was performed and significantly enriched GO Biological process and KEGG Pathways were identified (file10-11)(Fig3 D-E) .
Fig. 3 Crosstalk genes and functional enrichment analysis. (A) Venn diagram of differentially expressed genes obtained from AD and PD (B-c) Heat maps showing cross talk genes in AD and PD (D) Top 20 enriched biological processes in the crosstalk genes; (E) KEGG pathways significantly enriched in the crosstalk genes
We learnt that the crosstalk genes mainly regulated leukocyte migration, positive regulation of response to external stimulus, leukocyte differentiation, neutrophil activation, and other biological processes (Fig. 3D); participated in Leukocyte transendothelial migration, B cell receptor signalling pathway, TNF signalling pathway, NF-kappa B signalling pathway and so on (Fig. 3E).

3.3 Crosstalk genes PPI network

We extracted PPI relationship pairs for the crosstalk genes (file12) and constructed a PPI network (Fig. 4), which showed 4870 nodes and 9657 edges.

![Crosstalk genes PPI network](image)

**Fig. 4** Crosstalk genes PPI network (30 hub nodes obtained from topological analysis are displayed)
The topological properties of the network were analyzed (file13) and the top 30 hub node genes (Table2) were identified based on the degree of gene connectivity and considered important in the protein interaction network relationship.

### Table 2: Topological characteristics of the top 30 gene nodes in the crosstalk-gene PPI network

<table>
<thead>
<tr>
<th>name</th>
<th>AD_PD</th>
<th>Degree</th>
<th>AverageShortestPathLength</th>
<th>BetweennessCentrality</th>
<th>ClosenessCentrality</th>
<th>ClusteringCoefficient</th>
<th>TopologicalCoefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCAM1</td>
<td>up_up</td>
<td>673</td>
<td>2.612737</td>
<td>0.150601</td>
<td>0.38274</td>
<td>0.002503</td>
<td>0.003982</td>
</tr>
<tr>
<td>ITGA4</td>
<td>up_up</td>
<td>528</td>
<td>2.611707</td>
<td>0.099121</td>
<td>0.382891</td>
<td>0.003996</td>
<td>0.004326</td>
</tr>
<tr>
<td>VIM</td>
<td>up_up</td>
<td>320</td>
<td>2.645095</td>
<td>0.105724</td>
<td>0.35551</td>
<td>0.005602</td>
<td>0.006497</td>
</tr>
<tr>
<td>SHC1</td>
<td>up_up</td>
<td>273</td>
<td>2.812861</td>
<td>0.065675</td>
<td>0.35551</td>
<td>0.006502</td>
<td>0.009555</td>
</tr>
<tr>
<td>LYN</td>
<td>up_up</td>
<td>243</td>
<td>2.731451</td>
<td>0.072094</td>
<td>0.366106</td>
<td>0.006462</td>
<td>0.007585</td>
</tr>
<tr>
<td>IL7R</td>
<td>up_up</td>
<td>199</td>
<td>2.86892</td>
<td>0.035555</td>
<td>0.348563</td>
<td>0.002335</td>
<td>0.010105</td>
</tr>
<tr>
<td>UBC</td>
<td></td>
<td>184</td>
<td>2.253504</td>
<td>0.340648</td>
<td>0.443753</td>
<td>0.005702</td>
<td>0.010612</td>
</tr>
<tr>
<td>TXN</td>
<td>down_down</td>
<td>132</td>
<td>2.796785</td>
<td>0.027568</td>
<td>0.357553</td>
<td>0.0096</td>
<td>0.011983</td>
</tr>
<tr>
<td>TNFRSF1B</td>
<td>up_up</td>
<td>131</td>
<td>2.958368</td>
<td>0.032291</td>
<td>0.338024</td>
<td>0.01174</td>
<td>0.011297</td>
</tr>
<tr>
<td>TUBB2A</td>
<td>down_up</td>
<td>124</td>
<td>2.865829</td>
<td>0.030234</td>
<td>0.348939</td>
<td>0.003278</td>
<td>0.012209</td>
</tr>
<tr>
<td>CCND1</td>
<td>down_up</td>
<td>121</td>
<td>3.020816</td>
<td>0.02919</td>
<td>0.331036</td>
<td>0.00124</td>
<td>0.017982</td>
</tr>
<tr>
<td>BTK</td>
<td>up_up</td>
<td>112</td>
<td>2.934048</td>
<td>0.022542</td>
<td>0.340826</td>
<td>0.015766</td>
<td>0.017501</td>
</tr>
<tr>
<td>AIM2</td>
<td>up_up</td>
<td>109</td>
<td>3.703636</td>
<td>0.02737</td>
<td>0.296704</td>
<td>0</td>
<td>0.016514</td>
</tr>
<tr>
<td>PTP4A3</td>
<td>up_up</td>
<td>109</td>
<td>3.336974</td>
<td>0.024112</td>
<td>0.299673</td>
<td>1.70E-04</td>
<td>0.017308</td>
</tr>
<tr>
<td>MAPK13</td>
<td>down_up</td>
<td>103</td>
<td>3.092745</td>
<td>0.013791</td>
<td>0.323337</td>
<td>0</td>
<td>0.024619</td>
</tr>
<tr>
<td>HCK</td>
<td>up_up</td>
<td>100</td>
<td>2.968879</td>
<td>0.018268</td>
<td>0.336827</td>
<td>0.012525</td>
<td>0.020322</td>
</tr>
<tr>
<td>SERPINH1</td>
<td>up_up</td>
<td>96</td>
<td>2.955894</td>
<td>0.020851</td>
<td>0.338307</td>
<td>0.003728</td>
<td>0.017841</td>
</tr>
<tr>
<td>VAV1</td>
<td>up_up</td>
<td>95</td>
<td>2.930338</td>
<td>0.014357</td>
<td>0.341258</td>
<td>0.028219</td>
<td>0.019849</td>
</tr>
<tr>
<td>GNAI1</td>
<td>down_down</td>
<td>93</td>
<td>3.09357</td>
<td>0.029624</td>
<td>0.323251</td>
<td>0.001636</td>
<td>0.015197</td>
</tr>
<tr>
<td>WNK1</td>
<td>down_down</td>
<td>92</td>
<td>3.109439</td>
<td>0.022769</td>
<td>0.321601</td>
<td>4.78E-04</td>
<td>0.022952</td>
</tr>
<tr>
<td>UCHL3</td>
<td>down_down</td>
<td>90</td>
<td>3.072135</td>
<td>0.018117</td>
<td>0.325507</td>
<td>0.002747</td>
<td>0.021144</td>
</tr>
<tr>
<td>HYOU1</td>
<td>down_up</td>
<td>80</td>
<td>3.079555</td>
<td>0.015011</td>
<td>0.324722</td>
<td>0.001266</td>
<td>0.02724</td>
</tr>
<tr>
<td>RPN2</td>
<td>up_up</td>
<td>79</td>
<td>3.097486</td>
<td>0.01577</td>
<td>0.322843</td>
<td>0</td>
<td>0.027241</td>
</tr>
<tr>
<td>TXNDC5</td>
<td>up_up</td>
<td>78</td>
<td>3.068219</td>
<td>0.019356</td>
<td>0.325922</td>
<td>0</td>
<td>0.025968</td>
</tr>
<tr>
<td>RABAC1</td>
<td>down_up</td>
<td>71</td>
<td>3.022465</td>
<td>0.01991</td>
<td>0.330856</td>
<td>0.002414</td>
<td>0.016747</td>
</tr>
<tr>
<td>APOE</td>
<td>up_up</td>
<td>70</td>
<td>2.906018</td>
<td>0.020983</td>
<td>0.344113</td>
<td>0.004555</td>
<td>0.016556</td>
</tr>
<tr>
<td>PLCG2</td>
<td>up_up</td>
<td>69</td>
<td>2.955482</td>
<td>0.010744</td>
<td>0.338354</td>
<td>0.047315</td>
<td>0.024349</td>
</tr>
<tr>
<td>KRT10</td>
<td>down_up</td>
<td>68</td>
<td>2.854699</td>
<td>0.00891</td>
<td>0.3503</td>
<td>0.04302</td>
<td>0.027171</td>
</tr>
</tbody>
</table>
VCAM1, ITGA4 and VIM were noted as the top genes playing an important role in the network, and were upregulated in both AD and PD.

3.4 Crosstalk genes’ immune cell enrichment analysis

Using the xCell package, scores of immune-infiltrating cells corresponding to 364 samples of Crosstalk genes were calculated, ‘transform scores’ and ‘spillOver’ were applied to obtain the final corrected immune cell scores (file14-15). Scores of 55 immune cell types in the AD and PD datasets were noted and a heatmap was used to display the scores of immune infiltrating cells in AD and PD datasets (Fig. 5a).

A violin diagram drawn using ‘vioplot’ was used to show the performance of the scores of each immune infiltrating cell in both diseases (Fig. 5b-d). The difference in scores of immune infiltrating cells of Case samples of AD and PD datasets was tested using Wilcoxon’s test. The relationship between each immune cell type in the two diseases was tested, where the cells were grouped in three categories (Fig. 5b-d) (file16). We can see that several immune cell types are closely related in both diseases.
Basophils, CD4+ memory T−cells, CD4+ T−cells, CMP, Macrophages M2, MSC were significantly overexpressed in both AD and PD. To test the correlation between immune cells, a correlation analysis of xCell scores for Case samples for each immune cell type in the two data sets was applied (file17 file18) (Fig. 6 a b).

For AD, immune cell CD8+ Tem and Th2 cells were highly positively correlated (COR = 0.8196), CD8+ Tcm and CD8+ Tem were highly positively correlated (COR = 0.8127), CD8+ T- cells and CD8+ Tcm were highly positively correlated (COR = 0.7822). B-cells were highly negatively correlated with Basophils (COR = -0.7763). For PD, CD8+ Tcm and CD8+ Tem were highly positively correlated
0.9695), Tgd cells and Th2 cells were highly positively correlated (cor= 0.9657), CD8+ Tem and Tgd cells were highly positively correlated (0.9500), CD8+ Tem and Th2 cells were highly positively correlated (cor= 0.8954). MSC and Preadipocytes were highly negatively correlated (cor= -0.7754), CD4+ memory T-cells and Pericytes were highly negatively correlated (cor= -0.7173), Memory B-cells and pro B-cells were highly negatively correlated (cor= -0.6979).

3.5 Analysis of immune-related genes and immune cells

The immune-related genes downloaded from the literature included 15 Adaptive immune cells and 13 Innate immune cell types (Fig. 7a)(file19). The immune-related genes obtained from InnateDB and ImmPort databases are shown in File 20-21 (File 20 is the downloaded data, it contains human and mouse genes, we used human genes). We merged the immune-related genes acquired from literature, InnateDB, and ImmPort databases to obtain 3046 immune genes as the final immune-related gene dataset (file22).

We extracted the expression values of these 3046 immune genes in the Case samples of AD and PD and found 1142 immune genes were expressed in AD whereas 2396 immune genes were expressed in PD (file23-24).

Using the xCell algorithm, we obtained the expression scores of immune cells in the Case samples of AD and PD. Since the names of 64 cell types included in xCell were different from the names of 28 cell types listed in literature, we identified and listed the cell types (Table 3).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>aDC</td>
<td>iDC</td>
<td>Activated B cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>Adipocytes</td>
<td>Keratinocytes</td>
<td>Activated CD4 T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>Astrocytes</td>
<td>Ivy Endothelial cells</td>
<td>Activated CD8 T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>B-cells</td>
<td>Macrophages</td>
<td>Central memory CD4 T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>Basophils</td>
<td>Macrophages M1</td>
<td>Central memory CD8 T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD4+ memory T-cells</td>
<td>Macrophages M2</td>
<td>Effector memeoory CD4 T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD4+ naive T-cells</td>
<td>Mast cells</td>
<td>Effector memeoory CD8 T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD4+ T-cells</td>
<td>Megakaryocytes</td>
<td>Gamma delta T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD4+ Tcm</td>
<td>Melanoocytes</td>
<td>Immature B cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD4+ Tem</td>
<td>Memory B-cells</td>
<td>Memory B cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD8+ naive T-cells</td>
<td>MEP</td>
<td>Regulatory T cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----</td>
<td>------------------</td>
<td>----------</td>
</tr>
<tr>
<td>CD8+ T-cells</td>
<td>Mesangial cells</td>
<td>T follicular helper cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD8+ Tcm</td>
<td>Monocytes</td>
<td>Type 1 T helper cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>CD8+ Tem</td>
<td>MPP</td>
<td>Type 17 T helper cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>cDC</td>
<td>MSC</td>
<td>Type 2 T helper cell</td>
<td>Adaptive</td>
</tr>
<tr>
<td>Chondrocytes</td>
<td>mv Endothelial cells</td>
<td>Activated dendritic cell</td>
<td>Innate</td>
</tr>
<tr>
<td>Class-switched memory B-cells</td>
<td>Myocytes</td>
<td>CD56bright natural killer cell</td>
<td>Innate</td>
</tr>
<tr>
<td>CLP</td>
<td>naive B-cells</td>
<td>CD56dim natural killer cell</td>
<td>Innate</td>
</tr>
<tr>
<td>CMP</td>
<td>Neurons</td>
<td>Eosinophil</td>
<td>Innate</td>
</tr>
<tr>
<td>DC</td>
<td>Neutrophils</td>
<td>Immature dendritic cell</td>
<td>Innate</td>
</tr>
<tr>
<td>Endothelial cells</td>
<td>NK cells</td>
<td>Macrophage</td>
<td>Innate</td>
</tr>
<tr>
<td>Eosinophils</td>
<td>NKT</td>
<td>Mast cell</td>
<td>Innate</td>
</tr>
<tr>
<td>Epithelial cells</td>
<td>Osteoblast</td>
<td>MDSC</td>
<td>Innate</td>
</tr>
<tr>
<td>Erythrocytes</td>
<td>pDC</td>
<td>Monocyte</td>
<td>Innate</td>
</tr>
<tr>
<td>Fibroblasts</td>
<td>Pericytes</td>
<td>Natural killer cell</td>
<td>Innate</td>
</tr>
<tr>
<td>GMP</td>
<td>Plasma cells</td>
<td>Natural killer T cell</td>
<td>Innate</td>
</tr>
<tr>
<td>Hepatocytes</td>
<td>Platelets</td>
<td>Neutrophil</td>
<td>Innate</td>
</tr>
<tr>
<td>HSC</td>
<td>Preadipocytes</td>
<td>Plasma/dendritic cell</td>
<td>Innate</td>
</tr>
<tr>
<td>Smooth muscle</td>
<td>pro B-cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tgd cells</td>
<td>Sebocytes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Th1 cells</td>
<td>Skeletal muscle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Th2 cells</td>
<td>Tregs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10 adaptive and 10 innate immune cell types were noted in xCell and are marked with different colors. These 20 cell types were extracted for subsequent analysis (file25) and their scores were analysed (file 26-27). The fraction of these immune cells is depicted in Fig. 7 b.
CD4+ naive T-cells and B-cells were highly expressed in PD samples compared with AD samples. The expression of Eosinophils was higher in AD disease samples than in PD disease samples. Macrophages M2, Natural killer T cell (NKT) and CD8+ naive T-cells were highly expressed in AD and PD samples. Macrophages M2 and Natural killer T cell (NKT) are innate immune cells, while CD8+ naive T-cells are adaptive.

A violin diagram was drawn to depict the performance of scores of each immune infiltrating cell in both diseases (Fig. 8a) and differences were tested using Wilcoxon’s Test (file28) to test significant relationships of immune cells with the two diseases (P <0.05). Macrophages M2, Natural killer T cell
(NKT) and CD8+ naive T-cells were found to be significantly different in AD and PD. Then, we examined the correlation between the 20 immune cells in AD and PD (fig8 bc)(file29-30).

Fig. 8 (A) The expression levels of 20 immune cells in AD and PD; (B) Correlation of immune cells in AD; (C) Correlation of immunity in PD

CD8+ T-cells and CD8+ naive T-cells, CD4+ memory T-cells and CD4+ T-cells were highly positively correlated in AD and PD (table 4).

Table 4 20 types of immune cells highly positively correlated (COR >=0.6) in AD and PD

<table>
<thead>
<tr>
<th>Cell1</th>
<th>Cell2</th>
<th>PD_cor</th>
<th>PD_pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD8+ naive T-cells</td>
<td>CD8+ T-cells</td>
<td>0.81793</td>
<td>8.186E-105</td>
</tr>
<tr>
<td>CD4+ memory T-cells</td>
<td>CD4+ T-cells</td>
<td>0.801068</td>
<td>1.9407E-97</td>
</tr>
</tbody>
</table>
3.6 Consensus cluster analysis of immune cells based on immune genes

1142 immune genes found expressed in AD and 2396 immune genes found expressed in PD were subjected to Consensus Clustering. The maxK values were determined using average silhouette width, gap statistic, and the elbow method to find the optimal number of clusters for the AD and PD expression matrices (Fig. 9 a-f).
Fig. 9  Cluster number analysis using different methods  (A-C)  Average silhouette width, Gap statistic, and Elbow method to analyse the number of AD clusters;  (D-F)  Average silhouette width, Gap statistic, and Elbow method to analyse the number of PD clusters
As seen in Fig 9, the number of optimal clusters denoted by the three methods were different, which may be related to the large number of gene features and the differences in the algorithms. In AD, the maximum number of clusters was 5 and the minimum number was 2. In PD, the maximum number was 6 and the minimum number was 1. Clustering consistency results for 2-5 clusters in AD, and 2-6 clusters in PD were analysed. Key clustering consistency results for AD and PD are depicted in Fig. 10 a-f.
Fig. 10  Cluster consistency analysis. (A-B) Consistent cumulative distribution function (CDF) plots for AD and PD. This figure shows the cumulative distribution function of scores with different values of K, which is used to determine the approximate maximum value of CDF for a selected k value, and the cluster analysis result that is the most reliable. That is, the k
value with a small descending slope of CDF is considered.  

(C-D) Delta Area Plot of AD and PD: This figure shows the relative changes of areas under the CDF curve compared to k and k-1. When k=2, since there is no k=1, the first point represents the total area under the CDF curve at k=2 (that is, the area of the centre line in Figure AB), rather than the relative change in area.  

(E-F) Consistent clustering diagram of AD and PD

The results show that to get the final k value, the descending slope of the Central Line and the relative change of the area under the CDF curve between K and K-1 should be as small as possible. We finally choose K=4 for AD and PD both. Figure EF shows the correlation between AD and PD samples at the selected k values. The rows and columns of the matrix represent the samples. Consistency matrix values are shown in white to dark blue on a scale from 0 (impossible to cluster together) to 1 (always cluster together). The consistency matrix is arranged according to the consistency categories (tree at the top of the heatmap). The bar between the tree and the heat map is the category. The more scattered the dark blue squares, the weaker the clustering results. The cluster-consensus and item-consensus for AD and PD was analysed using the calcICL method in the ConsensusClusterPlus package (Fig. 11 a-d).
Fig. 11 Cluster-consensus and item-consensus for AD and PD. (A-B) Cluster-Consensus Plot for AD and PD. These figures show the cluster-consensus value of each cluster under AD and PD (The mean value of pairwise consensus values of members in the cluster). The higher the value, the higher the stability. It can be used to assess the cluster-consensus values under the same and between different k values. We can see in Figure A, for AD, when k=4, the mean values are high. In figure B, for PD, when k=4, the mean values are all also very high. (C-D) Item-consensus Plot for AD and PD: This figure shows the score of each sample for AD and PD when k=4. The vertical axis represents item-consensus values. Each sample has an item-consensus value corresponding to a different cluster. The vertical bar represents each sample, and the height of the vertical bar represents the total item-Consensus values of that sample. Above each sample is a vertical bar, and the colour of the vertical bar indicates which cluster the sample is divided into.

From Fig. 11 we can see whether the classification of each sample has sufficient fidelity, to help determine the k value. As shown in Fig. 10 and Fig. 11 we clustered the AD and PD disease samples in 4 clusters each (file 31, file 32). The sample clustering results and the 20 Adaptive and Innate immune
cell scores for AD and PD across all samples are shown (Fig. 7b). Next, we combined the immune cell fractions and sample cluster results for the Case samples from AD and PD (file 33-36) for subsequent analysis.

3.7 Adaptive and Innate Immune Cell bias analysis

For each consensus cluster, we calculated the cluster distribution of xCell scores of Adaptive and Innate immune genes in AD and PD and presented these in a box plot (Fig. 12a-d). The Kruskal-Wallis test was performed to test differences in scores of immune cells in different clusters.
In Fig. 12 we can see that there were significant differences between immune cells in different clusters, and the greater this difference, the more marked the difference between the clusters. We can also see the overall scores of immune cells in different clusters. The top 3 immune cells from the significant clusters in AD and PD were considered high expression immune cells that play an important role in disease pathology (Table 5).

**Table 5** Significant immune cell populations in AD and PD

<table>
<thead>
<tr>
<th></th>
<th>Adaptive Immune Cells Top 3</th>
<th>Innate Immune Cells Top 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cluster1</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td></td>
<td>B-cells</td>
<td>CD8+ T-cells</td>
</tr>
<tr>
<td>cluster2</td>
<td>B-cells</td>
<td>CD8+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td>cluster3</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td>cluster4</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td><strong>PD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cluster1</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td>cluster2</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td>cluster3</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
<tr>
<td>cluster4</td>
<td>B-cells</td>
<td>CD4+ memory T-cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD8+ naive T-cells</td>
</tr>
</tbody>
</table>

The results showed that B-cells, CD4+ memory T-cells and CD8+ naive T-cells were adaptive immune cells that were highly expressed in all 4 clusters of both diseases, and innate immune cells Macrophages M2 and NKT were similarly highly expressed in all clusters. Adaptive immune cells CD8+ naive T-cells were significantly different between cluster 2 and cluster 3 in AD and in PD, both.

We extracted the genes related to the immune-related genes from the CrossTalkGene dataset and obtained 112 genes in total. Then we extracted the expression values of these 112 genes in the Case samples of AD and PD. Correlation analysis was conducted by combining these values with the xCell scores. These two datasets pertaining to each cluster were subjected to correlation analysis. Correlation results were obtained for each of the 8 clusters and are depicted (Fig. 13 A-B) allowing an estimation of
immune cells bias in the different clusters, and also allow for selection of specific cluster of subjects for longitudinal study.

![Image of correlation analysis](image)

**Fig. 13** Correlation analysis applied to different clusters. (A) The correlation of adaptive immune-related genes in different clusters in AD and PD; (B) The correlation of innate immune-related genes in different clusters in AD and PD.

The results show immune cell biases in different clusters. For adaptive immune-related genes, cluster1 and cluster4 in PD were highly correlated with a variety of immune cells. In cluster2 and cluster3, immune genes were positively correlated with immune cells. Adaptive immune related genes were highly correlated with a variety of immune cells in cluster 4 of AD, and participated in a variety of immune patterns (Fig. 13 A). For innate immune related genes, multiple immune cells in cluster4 of PD and
cluster4 of AD were highly correlated, suggesting that innate genes were more active in cluster4 samples of PD and AD (Fig. 13 B).

3.8 Identification of Potential Biomarkers Using LASSO Logistic Regression

Datasets of expression values for Case and Control samples of AD and PD were obtained (file1-2). The expression values of the crosstalk genes were subjected to LASSO Logistic Regression to screen the crosstalk genes (Fig. 14 A-B). The intersecting genes were selected and 127 genes were recorded as the biomarker crosstalk gene (file37).

For Adaptive immune-related genes, we extracted 431 genes’ expression profiles, which included 210 genes found in AD and 408 genes found in PD datasets and applied LASSO Logistic Regression (Fig. 14c-d). The intersecting genes among AD and PD were selected and a total of 78 genes were recorded as the biomarker adaptive immune genes (file38).

For Innate immune genes, we combined the innate immune genes obtained from literature with those obtained from InnateDB data to obtain 1335 Innate immune genes. 571 such genes were found in AD and 1183 in PD. LASSO Logistic Regression was applied (Fig. 14 E-F) and a total of 32 intersecting genes were recorded as biomarker Innate immune gene (file39). A diagram displayed the variation of the remaining variables' gene coefficients with different lambda values from the LASSO regression analysis.
Fig. 14  The optimal lambda values obtained from LASSO regression modelling of innate immune genes and adaptive immune genes. (A-B) Crosstalk genes: relationship between lambda value and mean square error in AD and PD Lasso regression analysis. The abscissa is log(lambda) and the ordinate is mean square error. There are two dashed lines in the figure, one is the
value of \( \lambda \) with the minimum mean square error and the other is the value of \( \lambda \) with the standard error from the minimum mean square error. (C-D) Adaptive immune genes: lambda value and mean square error in AD and PD Lasso regression analysis. (E-F) Innate immune genes: lambda value and mean square error in AD and PD Lasso regression analysis.

We obtained 3 genes from the intersection of the biomarker crosstalk genes and the biomarker adaptive immune genes, and 1 gene (DEFB1) from the intersection of the biomarker crosstalk genes with the biomarker innate-immune genes. ROC analysis using the expression values of these 4 genes yielded 3 genes (DUSP14, F13A1, SELE)(Fig. 15 A-B).
Fig. 15  Predictive efficacy of DUSP14, F13A1, SELE and DEFB1 in AD and PD. (A-B) ROC results of DUSP14, F13A1, SELE and DEFB1 in AD and PD; (C) Functional correlation analysis results of DUSP14, F13A1, SELE and DEFB1

AUC (AUC>70%) values obtained for DUSP14, F13A1, SELE and DEFB1 in discriminating PD were higher than those for AD prediction. SELE performed better than the other 3 genes in discriminating both AD and PD. To further analyse the functions of these genes, we obtained datasets of human KEGG pathways and related genes and mined the corresponding pathways and then isolated all the genes in
each pathway. We examined whether there is interaction between each pathway and the biomarker
crosstalk genes, biomarker adaptive immune genes, and biomarker innate immune genes (Fig 15C)
(file40). The results showed that SELE, an adaptive immune gene, mainly regulates the TNF signalling
pathway, cell adhesion molecules (CAMs) and fluid shear stress and atherosclerosis. Within the TNF
signalling pathway, VCAM1 was also involved in Cell adhesion Molecules (CAMs). Within Cell
adhesion Molecules (CAMs), the adaptive immune gene ITGB2 regulates both Staphylococcus aureus
infection and complement and coagulation cascades. From Fig. 15c, we can see that F13A1 is mainly
involved in the regulation of complement and coagulation cascades pathway. DEFB1, an innate immune-
related gene, is mainly involved in the regulation of Staphylococcus aureus infection and ABC
transporters. Within Staphylococcus aureus infection, other crosstalk genes such as KRT24 and FCGR2B
also participate in the regulation. In addition, ITGB2 and other genes are associated with other pathways
to regulate the immune function in both AD and PD. It can be inferred from the above that immune-
related crosstalk genes interact with other genes and jointly influence the two diseases.

4 Discussion

The present bioinformatic study applied immunocorrelation analysis to identify immune-related genes,
cells and pathways that might serve as key linkage mechanisms between AD and PD. We found that
innate immune cells M2 macrophages and NKT are highly expressed in both AD and PD. M2
macrophages are primarily involved in the Th2 immune response. Th2 cells produce cytokines that
promote the humoral immune response, including IL-4, IL-5, IL-6, IL-10 and IL-13[29]. NKT cells
mediate proinflammatory and immunomodulatory effects, which range from B-cell regulation,
production of specific antibodies, suppression of autoimmunity to cytokine production, dendritic cell
crosstalk, and T/B cell interactions[30]. Among the adaptive immune cells, B-cells, CD4+ memory T-
cells and CD8+ naive T-cells were found highly expressed in all 4 clusters of AD and PD. B cells might
exert protective functions in periodontitis. B-cell-deficient mice show alveolar bone loss without
bacterial infection, while clinical evidence shows that B cells and plasma cells, along with
osteoclastogenic factors, are involved in alveolar bone destruction in periodontitis[31].

Using a data mining approach with a series of reductive analyses, we obtained a 3 gene set, DUSP14,
F13A1 and SELE, as key crosstalk genes linking PD and AD, which was largely supported by
experimental and clinical data. The mechanistic role of DUSP14 in mediating AD and PD is not investigated. There are few reports regarding the roles of DUSP14 in the literature, mainly focusing on pathways related to T cells. DUSP14 can downregulate T-cell receptor signalling by inhibiting TGF-β-activated kinase 1-binding protein 1 (TAB1) activation[32]. DUSP14 is a mitogen-activated protein kinase phosphatase that plays a critical role in the regulation of T cell activity. TRAF2 mediated Lys63-linked ubiquitination of DUSP14 leads to DUSP14 activation in T cells[33]. DUSP14 directly interacts with TGF-beta-activated kinase 1 (TAK1)-binding protein 1 (TAB1) and dephosphorylates TAB1 at Ser(438), leading to TAB1-TAK1 complex inactivation in T cells and can downregulate T-cell receptor (TCR) signalling by inhibiting TAB1 activation[32]. Activated DUSP14 also directly dephosphorylates extracellular signal-regulated kinases (ERK) and attenuates the ERK signalling pathway. TRAF2-mediated ubiquitination of Lys63-linked DUSP14 also enhances its phosphatase activity[34]. Protein arginine methyltransferase (PRMT)5-mediated arginine methylation may sequentially stimulate TRAF2-mediated DUSP14 ubiquitination and phosphatase activity, leading to inhibition of TCR signalling[35]. Therefore, enhancement/activation of DUSP14 or DUSP14 upstream molecules is a potential modality for the attenuation of autoimmune diseases such as systemic lupus erythematosus (SLE)[36]. The role of F13A1 has been investigated in AD. F13A1 subunit was detected by immunohistochemistry in a subset of AD reactive microglia, while F13A1 Val34Leu gene polymorphism is associated with sporadic AD where homozygous LL genotype shows about a fourfold higher risk of developing AD compared to the homozygous VV genotype[37]. F13A1 may also influence the maintenance of neural connections[38]. The F13A1 204Phe allele is strongly associated with ischemic stroke in young women and the homozygous genotype (Phe/Phe) are associated with manyfold higher stroke risk than heterozygous (Tyr/Phe) genotype[39]. Functionally, a pro-angiogenic function of F13A1 is affected by the interaction between vascular endothelial growth factor receptor 2 (VEGFR2) and integrin αvβ3 on the cell membrane, which facilitates important steps in granulation tissue formation at wound sites. F13A1 deficiency can thus present as intracranial haemorrhage, delayed bleeding or chronic wounding of the skin and impaired mucosal healing. F13A1 thus functions to link primary hemostasis, coagulation, and definitive tissue healing. Another important recently identified function of F13A1 is its ability to control cellular infiltration by binding to specific macromolecules, thereby limiting bacterial spread at the wound site and promoting host cell migration and survival[40]. In the brain, F13A1 expression has been detected by immunohistochemistry in reactive microglia during glioma formation, which is a distinctive feature
of AD pathogenesis[37]. F13A1 levels are gradually elevated from controls to mild cognitive impairment (MCI) and AD. More importantly, F13A1 in the serum proteome can serve as a potential non-invasive early diagnostic marker of MCI and AD[41]. SELE encodes for E-selectin and is involved in Leukocyte/endothelial cell adhesion, and its expression is reported to increase 4-fold in Treponema denticola oral infections[42], a subgingival oral spirochete species which is a key periodontal pathogen[43]. Its role in several age-associated conditions such as age-related macular degeneration[44] and other conditions. SELE has been found to be related with peripheral arterial occlusive disease[45]. The serum level of SELE has been found significantly elevated in systemic sclerosis with early onset disease [46]. Regarding periodontitis, the Ser128Arg polymorphism is associated with periodontitis [47]. SELE expression is also found positively correlated with the duration of Sjögren's syndrome, characterised by dysregulation of circulating immune cells, T cells and antigen presenting cells and vascular endothelial extravasation[48, 49]. In an animal model of AD, SELE expression was found significantly elevated, indicating its role in AD development[50]. The cell-surface glycoprotein E selectin plays an important role in immune adhesion[51]. It is also associated with the accumulation of white blood cells at sites of inflammation by mediating cell adhesion to the intima of blood vessels. As a clinical diagnosis, AD shows variable pathology. Clinically, E-selectin has been found significantly raised in the cerebrospinal fluid (CSF) of AD patients without typical signature biomarker profiles, suggesting it may specifically mark the vascular mechanisms underlying AD pathology[52]. The SELE Ser128Arg gene polymorphism has also been associated with AD[53, 54, 55] and SELE polymorphisms are also associated with Lewy body dementias [56].

Overall, the findings of this bioinformatic study were supported by existing experimental evidence addressing PD and AD but the roles of the discovered biomarkers DUSP14, F13A1 and SELE in mediating the link between the two diseases has not been addressed. The present data must be considered as a theoretical premise for further investigation that explores the validity of these biomarkers in large-scale clinical trials and their mechanistic roles in experimental or translational research focused on immune mechanisms implicated in AD and PD linkage.

5 Conclusion
Bioinformatic analysis integrating experimental transcriptomic data from Alzheimer’s disease and periodontitis revealed the most robust potentially shared molecular linkages. Three biomarker crosstalk genes; DUSP14, F13A1 and SELE were identified as the most robust signature. Macrophages M2 and NKT among innate immune cells, and B-cells, CD4+ memory T-cells and CD8+ naive T-cells among adaptive immune cells emerged as top immune cells linking PD and AD. These findings warrant future research in experimental and clinical studies.

Reference


[10] Puris, E., S. Auriola, P. Korhonen, S. Loppi, K.M. Kanninen, T. Malm, J. Koistinaho, and M. Gynther, Systemic Inflammation Induced Changes in Protein Expression of ABC Transporters and Ionotropic Glutamate Receptor Subunit 1 in the Cerebral Cortex of Familial Alzheimer’s


**Declarations**

- **Competing interests**  The authors declare no potential conflict of interest in this paper.
- **Funding**  This research received no external funding.
- **Conflict of interest**  The authors declare no conflict of interest.
- **Ethics approval**  As this study only applied bioinformatic techniques based on computational analyses of publicly available datasets, therefore, this study did not require ethical approval.
- **Consent to participate**  Not applicable
- **Consent for publication**  Not applicable
- **Code availability**  Not applicable

**Availability of data and materials**

All the raw data of this study were obtained from public open access. We downloaded expression profile datasets related to periodontitis (PD) and Alzheimer’s disease (AD) from the GEO (https://www.ncbi.nlm.nih.gov/geo/) database. The datasets are listed in table1. We downloaded protein-protein related gene pairs from MINT (http://mint.bio.uniroma2.it/mint/Welcome.do), HPRD (http://www.hprd.org/index_html), BIOGRID (http://thebiogrid.org/), DIP (http://dip.doc-mbi.ucla.edu/dip/Main.cgi), mentha (http://mentha.uniroma2.it/index.php), PINA (http://cbg.garvan.unsw.edu.au/pina/), InnateDB (http://www.innatedb.com/), and Instruct (http://instruct.yulab.org/index.html) databases. The data used to support the findings of this study are also available from the corresponding author upon request.

**Authors’ contributions**

Jieqi Jin (Email: jjq007@126.com) and Mengkai Guang (Email: tygmk@qq.com) contributed equally as the first author. Gerhard Schmalz (Email: gerhard.schmalz@medizin.uni-leipzig.de) contributed as the senior author. Xiaofeng Huang (Email: huangxf1998@163.com) contributed as the corresponding author.

Conceptualization, Xiaofeng Huang and Gerhard Schmalz; methodology, Jieqi Jin; software, Mengkai Guang, Xing Yan and Menglin Cheng; validation, Jieqi Jin, Yong Liu and Liwei Zhang; formal analysis, Bo Zhang; investigation, Menglin Cheng; resources, Jieqi Jin and Mengkai Guang; data curation, Jieqi Jin and Mengkai Guang; writing—original draft preparation, Jieqi Jin and Mengkai Guang; writing—review and editing, Xiaofeng Huang and Gerhard Schmalz; visualization, Mengkai Guang; supervision, Xiaofeng Huang; project administration, Xiaofeng Huang.

**Acknowledgements:**

I want to thank my colleagues, for their collaboration and help during gathering data for my research project. I would also thank my parents and my lovely kid for their kindness to let me have time and energy to finish this work.

**Authors’ information**

First author: Jieqi Jin
Date of birth: 1989.06.10 in China
Email: jjq007@126.com
Affiliation: Department of Stomatology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China
EDUCATION
2015-2018  Doctoral degree
University of Munich, Germany. Supervised by Prof. Dr. Dr. Matthias Folwaczny. (Conservative Dentistry and Periodontology)
2007-2015  Master & Bachelor’s degree
West China School of Stomatology, Sichuan University, China. Supervised by Prof. Hongkun Wu.

RESEARCH TOPICS AND SKILLS
Oral Health & Medicine
Periodontology
Oral and Maxillofacial Medical imaging
Oral and systemic interactions
Bioinformatic studies
Preventive Medicine

First author: Mengkai Guang
Family Name: Guang
Given Name: Mengkai
Date of Birth: Aug, 30, 1990
Gender: Male
Nationality: China
Email: tygmk@qq.com, gmk0351@gmail.com
Address: No.2 Yinghuadong St. Chaoyang Qv, Beijing, 100029, China

Education:
2014-2017 Double Degree PhD. student in Dental Pharmacology, Graduate School of Dentistry, Tohoku University, Japan
2012-2017 Doctor student in Dental Implant Department, West China School of Stomatology, Sichuan University, China
2007-2012 Bachelor’s degree in Clinical Dentistry, West China School of Stomatology, Sichuan University, China

Clinical & Laboratory Expertise:
Solid academic knowledge of clinical dentistry
In vitro experiments (Primary rat calvaria osteoblast cell collection, Transwell co-culture and migration, Patch clamp, MTT, CCK8, Cell cycle, PCR, ELISA, Flow cytometry, etc.)
In vivo experiments (Rat, Mouse, Chick chorioallantoic membrane)

Corresponding author: Xiaofeng Huang DDS, PhD
Professor and Director, Department of Stomatology, Beijing Friendship Hospital Affiliated to Capital Medical University

Learning experience:
2002-2005, School of Stomatology, Peking University, PhD
1998-2000, School of Stomatology, Peking University, Master
1993-1998, School of Stomatology, Peking University, DDS

Work experience
2007-2009, Center for Craniofacial Molecular Biology, University of Southern California, United States of America, Postdoctoral fellow and visiting scholar
2000-2023 Department of Stomatology, Beijing Friendship Hospital Affiliated to Capital Medical University, China

Current and previous research interests:
Orthodontics; relationship between dentistry and systemic diseases; Developmental biology; Tissue regeneration; Bone metabolism
Memberships of professional societies and awards received:
2021, Education and Teaching Award, Capital Medical University
2019, Top papers of Chinese Journal of Stomatology
2018, Second Prize (3rd place), the China First Medical Innovation and Entrepreneurship Competition (Shenzhen)
2016, Excellent Instructor for Standardized Training of Resident, Beijing, China
2015, Outstanding Tutor award, Capital Medical University