The significance of lysosome in the diagnosis and subclassification of Alzheimer's disease

Chenghao Li
Yanbian University Hospital(Yanbin hospital)*

Zhuo Zheng
Yanbian University Hospital(Yanbin hospital)*

Meishan Jin (jinmeishan1210@126.com)
Yanbian University Hospital(Yanbin hospital)*

Research Article

Keywords: lysosome, Alzheimer's disease, heterogeneity, autophagy, immunity

Posted Date: November 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2305864/v1

License: This work is licensed under a Creative Commons Attribuition 4.0 International License.
Read Full License
Abstract

Background: Alzheimer's disease (AD) is a heterogeneous disease with complex pathophysiological characteristics. Lysosomes are the main organelles of degradation in eukaryotic cells, and their dysfunction is closely related to AD. Therefore, our goal is to identify the lysosomal induced molecular subtype of AD and further explore the possible mechanisms.

Methods: The dataset was downloaded from the GEO database. By differential expression analysis, 50 differentially expressed lysosomal genes in AD were identified. R-package "ROCR" was used to plot and calculate ROC curves and AUC values for differential lysosomal genes. The expression data of the above genes with AUC greater than 0.7 in the AD group were extracted and the R package "ConsensusClusterPlus" was used for consistent clustering of the AD data set. The contents of 28 kinds of immune cells in all samples (cluster1:28, cluster2:28) of the AD dataset were calculated using the R-package "GSVA". The R package "limma" was used to analyze the differences of autophagy genes in 56 AD data sets based on consistent clustering. R package "WGCNA" carried out weighted co-expression network analysis of the differential genes between cluster1 and cluster2, and extracted the related genes of the two modules with the highest positive and negative correlation. GO and KEGG functional enrichment of the above module genes was performed.

Results: Two lysosomes subtypes (Cluster1:28, Cluster2: 28) with different outcomes were identified in AD cohort by unsupervised clustering of lysosome diagnostics molecular, known as cluster 1/2. The results showed that 13 immune cells were significantly different between cluster1 and cluster2. A total of 76 differentially expressed autophagy genes were identified. It indicated autophagy heterogeneity between lysosome subtypes. The enrichment analysis of the key module genes of the lysosome subtype showed that the key module genes were mainly concentrated in the gene set associated with the synapses.

Conclusions: According to the heterogeneity of lysosomes, we identified two different AD subtypes based on different lysosome gene expressions, preliminarily revealing that the heterogeneity of AD may be mainly caused by lysosomes. The role of lysosomes may be related to autophagy and synapses.

1 Introduction

Alzheimer's disease (AD) is a progressive, persistent and degenerative disease of the central nervous system (CNS), and is a heterogeneous disease with complex pathophysiological characteristics[1]. The main clinical manifestations were impairment of recognition function and memory ability[2]. The main symptoms and neuropathologic criteria for the diagnosis were plaques formed by extracellular amyloid β (Aβ) deposition in the brain and hyperphosphorylated tau neurofibrillary tangle accumulation in nerve cells and loss of neurologic elements[3]. However, recent fundamental discoveries have highlighted the important pathological role of other key cellular and molecular processes[1]. Despite this, The pathogenesis of the disease is not clear, and there is a lack of effective treatment[4–5]. It's valueable to
identify different subtypes of AD and adopt appropriate treatment for patients with different subtypes, so as to improve treatment of AD.

With the aging of the world population, the incidence of AD is increasing. According to the 2021 International AD Report, the number of dementia patients in the world has exceeded 55 million by 2021, of which 50%–75% are AD patients. There are about 6.5 million Americans aged 65 and older with AD. It is estimated that the prevalence of AD in America will exceed 13.8 million by 2060\(^6\). Aβ is produced by the cleavage of APP (amyloid precursor protein) by two enzymes, β-secretase and γ-secretase, and lysosomes are considered as one of the main sites of cleavage\(^7\). The activity of β-secretase is around pH 4.5, which is the pH of lysosomes\(^8\). γ-secretase needs to form a complex with the AD-related protein Presenilin 1 (PS1) and several other proteins to exert its enzymatic activity\(^9\). Mutation of PS1 encoding gene is the main cause of familial AD, and PS1 mutation also affects autophagic flow and lysosomal enzyme activity\(^10\). If lysosomal function is impaired or Aβ production is increased, Aβ accumulates in neurons, leading to cell death and thus the pathological response to AD\(^11\). Therefore, it is very important to reveal the mechanism of lysosomal function in AD.

Lysosomes are the main organelles of degradation in eukaryotic cells, and their dysfunction is closely related to many diseases\(^12\). As the center of cell hydrolysis, lysosomes contain a large number of hydrolytic enzymes, which form autophagic lysosomes in the process of autophagy and participate in the degradation of autophagy substrates\(^13\). Amp-activated protein kinase (AMPK) and mammalian target of rapamycin complex 1 (mTORC1) are key proteins in autophagy regulation and play opposite roles in autophagy regulation when activated\(^14\). As a signal center, lysosomes are involved in cell energy sensing and amino acid sensing through AMPK and mTORC1, respectively\(^14–17\). Lysosomes are integrated in response to various signal stimuli to regulate the process of autophagy. Autophagy, as the main mechanism of maintaining cell homeostasis and prolonging biological life, also plays an extremely important role in AD\(^18\). Therefore, understanding the signal regulation of autophagy by lysosomes can provide new ideas for the treatment of AD targeting lysosomes. To ensure greater progress in AD treatment, it is important to consider that AD is a heterogeneous multifactorial disease. However, it is still not possible to accurately classify AD patients at the molecular biological level. According to the identification of different subtypes of AD at the molecular biological level, appropriate treatment options can be taken for patients with different subtypes. However, the details of lysosomal mediated AD heterogeneity remain unclear.

In this study, we analyzed the expression of lysosomal related genes in AD based on GEO database system. Firstly, we tried to screen lysosomal related biomarkers in AD by ROC curve analysis. In addition, we classified AD into different lysosomal types by unsupervised clustering based on lysosomal related biomarkers. In addition, to elucidate the reasons why lysosomal subtypes influence the different progression of AD, we performed WGCNA analysis, differential expression of autophagy genes between subtypes, and immune infiltration analysis.
2 Methods And Materials

2.1 Data sources and information

GSE122063 dataset was downloaded from GEO database, containing 136 samples (AD: 56, Control: 44, VaD: 36), AD and Control samples were used in this analysis. Details of the sample are provided in the supplementary materials GSE22063_sample_info. The R package "limma" was used for differential analysis of the dataset (|logFC|>0.5 & Pvalue < 0.05).

The lysosome-related gene set was downloaded from the GSEA MSigDB database (http://software.broadinstitute.org/gsea/msigdb).

The intersection of lysosome-related gene sets and differential genes was conducted.

2.2 Screening diagnosis-associated lysosome-related gene in Alzheimer's disease

Receiver operator characteristic curve (ROC) is a curve drawn by true positive rate and false positive rate, which can represent the relationship between sensitivity and specificity and is often used in diagnostic experiments. R package "ROCR" was used to draw and calculate the ROC curve and AUC value of intersection differential genes, and the genes with AUC > 0.7 were screened for subsequent analysis.

2.3 Consensus clustering

Consensus clustering is a method that provides quantitative evidence for determining the number and membership of possible clusters in a data set, for example as microarray gene expression. This approach has been widely used in cancer genomics, where new molecular subclasses of disease have been discovered. The R package "ConsensusClusterPlus" was used to cluster the expression data of 56 AD samples corresponding to the genes with AUC > 0.7 based on Euclidean distance and KM algorithm. To further verify the accuracy of the grouping, we used the R package "FactoMineR" to perform PCA analysis on the classification results.

2.4 Heterogeneity among lysosome subtypes

Autophagy related genes were downloaded from HHAD database (http://www.autophagy.lu/clustering/index.html). R package "limma" was used to analyze the differential expression of autophagy related genes in 56 AD datasets based on consensus clustering (|logFC|>0.5 & Pvalue < 0.05).

ssGSEA (single sample gene set enrichment analysis, ssGSEA) was conducted for immune infiltration analysis. 28 immune cell genes was retrieved from the database (http://cis.hku.hk/TISIDB/download.php). The R package "GSVA" was used to calculate the content of 28 immune cells in all samples (Cluster 1: 28, Cluster 2: 28) of the AD dataset. The R package "ggplot2" was used to draw box diagrams by wilcox.test.
2.5 Construction of co-expression network and enrichment analysis

The R package "limma" was used to perform differential expression analysis on the datasets clustered coherently (|logFC|>0.5 & Pvalue < 0.05). The R package "WGCNA" was used to perform weighted co-expression network analysis on the above obtained differential genes to extract the related genes of the two modules with the highest correlation.

R package "clusterProfiler" was used to GO and KEGG enrichment analysis of the above module genes.

3 Results

3.1 Lysosome gene set in Alzheimer's disease

To compare the transcriptional expression of the Lysosome genes between AD and Control samples, the lysosome gene set from GSEA MSigDB and RNA sequencing profiles from GEO databases were used. The results revealed that there were 5447 differentially expressed genes, including 2389 differentially up-regulated genes and 3133 differentially down-regulated genes (|logFC|>0.5&Pvalue < 0.05, Fig. 1A). The Fig. 1B shows the heat map of the top 10 up-and down-regulated genes. The lysosomal gene set (286 genes) was intersected with the differential genes. Fifty differentially expressed lysosomal genes were obtained, including 12 up-regulated genes and 38 down-regulated genes (Fig. 1C, Supplementary material: lysosome_deg).

3.2 The diagnosis associated lysosome gene in Alzheimer's disease

The R package "ROCR" was used to draw ROC curves for 50 differentially expressed lysosomal genes, and the corresponding AUC values were calculated to screen AUC > 0.7. This section retained 44 genes for further analysis. Only the ROC curves of up-regulated genes were shown in Fig. 2, and the down-regulated genes were shown in the supplementary materia:Diagnostic ROC curve.

3.3 Lysosome subtypes in Alzheimer's disease

Based on Euclidean distance, the expression data of 56 AD samples corresponding to the above 44 genes were clustered by KM algorithm. As can be seen in Fig. 3A&B, when k = 2, the classification result was the best. 56 samples were divided into two categories (Cluster1: 28, Cluster2: 28). To further verify the accuracy of the grouping, we used the R package "FactoMineR" to perform PCA analysis on the above classification results. As can be seen from Fig. 3, the distinction between the two groups is relatively high.

3.4 Immune status heterogeneity between lysosome subtypes
To further understanding the relationship between lysosome subtypes and Immune infiltration in AD, we applied the single sample gene set enrichment analysis, ssGSEA, to dissect the composition of diverse immune cells in two lysosome subtypes. As shown in Fig. 4, the relative abundances of 28 different immune cells have been characterized. The results showed that a total of 13 immune cells were significantly different between cluster1 and cluster2. Among them, the high concentrations in cluster1 were CD56dim Natural killer cell, Plasmacytoid dendritic cell, MDSC, Effector memeory CD8 T cell, Natural killer cell, Natural killer T cell, Neutrophil. The higher expression level in cluster2 were Immature dendritic cell, Effector memeory CD4 T cell, T follicular helper cell, Activated CD8 T cell, Type 2 T helper cell and Activated B cell.

3.5 Autopaghy heterogeneity between lysosome subtypes

We downloaded autophagy related genes (223) from HHAD database and analyzed the differential expression between two lysosomal subtypes to explore whether the autophagy in the AD patients is affected by the different lysosome subtypes. Among 223 autophagy genes, 219 genes had expression data in AD samples (Fig. 5A). A total of 76 differentially expressed genes were identified (|logFC|>0.5& P value < 0.05), among which 62 differentially up-regulated genes and 22 differentially down-regulated genes were identified (Fig. 5B-C).

3.6 Construct weighted co-expression network and enrichment analysis

A total of 2819 differentially expressed genes were identified between the two lysosomal subtypes, including 1808 differentially up-regulated genes and 1011 differentially down-regulated genes (Fig. 6). All the above 2819 differentially expressed genes were included in WGCNA analysis. 4 modules were identified by WGCNA in AD. First, we clustered the samples, and by observing the clustering situation of the samples, we checked whether there were outlier samples to be deleted, so as to ensure the accuracy of subsequent analysis. In Fig. 7A, there are no abnormal outlier samples, so there is no need to delete the samples. The power threshold is determined to be 16, that is, the ordinate R^2 in Fig. 7B (left) is around 0.8. It indicated that the network is closer to the scale-free distribution. The mean value of the adjacency function in Fig. 7B (right) is also gradually approaching 0, showing a gentle trend. Merge similar modules based on a shear height of 0.2 (80% similarity, Fig. 7C). Based on module clustering, we combined similar modules, resulting in four modules. We show the clustering tree of four modules (7D). Cluster 1 was positively correlated with grey module (including 102 genes), with correlation coefficient of 0.91 and P value as 9×10^{-22} (Fig. 7E). Cluster 2 was positively correlated with turquoise module (including 1,269 genes), with correlation coefficient of 0.88 and P value of 9×10^{-19} (Fig. 7E). Genes from these strongly correlated modules were defined as characteristic genes.

Genes from these strongly correlated modules were enriched based on GO and KEGG pathways to find the common functions and related pathways of a large number of genes in the gene set. 562 GO functional items were significantly enriched. The top 15 terms for GO-molecular function (MF), GO-biological process (BP), and GO-cell component (CC) was selected (Fig. 8A). The top 5 BP groups were
enriched, including regulation of membrane potential, synapse organization, modulation of chemical synaptic transmission, regulation of trans-synaptic signaling, synaptic vesicle cycle. The top 5 CC groups were enriched, including presynapse, synaptic membrane, transmembrane transporter complex, transporter complex, and postsynaptic membrane. The top 5 MF groups were enriched, including ion channel activity, metal ion transmembrane transporter activity, gated channel activity, voltage-gated ion channel activity, and voltage-gated channel activity. The top 5 KEGG pathways were enriched, including neuroactive ligand receptor interaction, calcium signaling pathway, cAMP signaling pathway, adrenergic signaling in cardiomyocytes, GABAergic synapse (Fig. 8B).

4 Discussion

AD is a heterogeneous disease with complex pathophysiological features. There is currently no method to accurately classify AD patients at the molecular biological level. Therefore, it is very important to accurately classify the different subtypes of AD at the molecular biological level. Different subtypes of AD patients can take the corresponding treatment plan. Current studies have shown that the expression of lysosomal related genes is closely related to the heterogeneity and diagnosis of AD. Two lysosomal gene subtypes, cluster 1 and cluster 2, were identified in AD based on the expression of most of the abnormally expressed lysosomal associated genes. The two lysosomal subtypes showed significantly different autophagy gene expression, immune status, and biological processes. To our knowledge, this is the first study to map the transcriptome extent of lysosomal genes, focusing on the landscape and function of lysosomal genes in AD.

A total of 44 lysosomal related genes were confirmed to be more important for the diagnosis of AD. Unlike the vast majority of cells in the body, neurons cannot dilute damaged organelles and cellular waste through cell division, but must maintain neuronal homeostasis through an efficient intracellular clearance system\textsuperscript{[19]}. Therefore, lysosomes are essential for neuronal function maintenance and survival. Many studies have suggested that the loss of lysosomal function is closely related to the occurrence and development of AD\textsuperscript{[20–22]}. The accumulation of harmful substances in cells caused by abnormal lysosomal function is an important reason for the occurrence and development of AD\textsuperscript{[23–24]}. The current study further revealed the diagnostic role of lysosome-related genes in AD. Future studies can further validate and investigate the fine mechanisms of these lysosome-related genes in AD. This study found that autophagy heterogeneity and immune heterogeneity in AD are closely related to lysosomal function. AD patients are divided into finer categories with different autophagy and immune status. With the deepening of research, more and more evidence shows that the pathogenesis of AD is not only limited to neurons, but also includes the interaction between the brain and the immune system\textsuperscript{[25]}. Continued formation and deposition of Aβ leads to activation of the innate immune system to form neuroinflammation. Experimental, genetic, and epidemiological data show that activation of the innate immune system plays a crucial role as a factor contributing to the progression of AD\textsuperscript{[26]}. One study found that immune cell subsets of the adaptive immune system may play a role in AD\textsuperscript{[27]}. The study isolated the immune cell subset TEMRA in the blood from healthy people and patients with AD as well as people with
mild cognitive impairment and found that these immune cell subsets can remember and respond to specific foreign invaders. This also indicates that these immune cell subsets will play a role in AD. Recent research has proposed a new theory, Alzheimer's disease is not a brain disease per se, but rather an autoimmune disease that occurs primarily in the brain\[^{28}\]. Autophagy is closely related to the occurrence of AD\[^{29–30}\]. Autophagy is a highly conserved process of eukaryotic cell disposal and recycling in the process of evolution\[^{31}\]. Under gene regulation, autophagy can maintain the stability of the body's internal environment by removing aging damaged organelles and misfolded proteins\[^{32}\]. When autophagy is impaired, on the one hand, autophagic vesicles (AVs) will decompose self-encapsulated APP to generate A\(_\beta\), in which \(\beta\)-lamellar structure is also transported to autophagy bodies to be degraded by \(\gamma\)-secretase to generate more A\(_\beta\)\[^{33}\]. However, the retained AVs cannot be degraded by lysosomes, leading to A\(_\beta\) clearance obstacle, namely, autophagy disorder. Abnormal accumulation of A\(_\beta\) can be caused by affecting the generation, clearance and secretion of A\(_\beta\). On the other hand, autophagy pathway is an important way for neurons to degrade Tau protein. Autophagy disorder will affect the clearance of abnormally phosphorylated Tau protein, resulting in abnormal accumulation of phosphorylated Tau protein, which in turn will affect the process of autophagy and further accelerate the pathological process of AD\[^{34}\]. These results provide a rationale for predicting clinical outcomes and meaningful research targets for AD.

Subsequent characterization of differentially expressed genes in lysosomal subtypes suggested that synapses play critical roles in distinct lysosomal subtypes. Normal aging leads to cognitive decline due to a decrease in synaptic contact maintenance and a decrease in the density of synaptic ridges. The size of dendritic ridges is related to their susceptibility to aging, with small ridges more susceptible to aging than larger ones. Synapse loss is an early event in AD development and is caused by soluble \(\beta\)-amyloid, accumulation of phosphorylated tau, and increased mitochondrial production of free radicals at synapses. Therefore, reducing the synaptic accumulation of amyloid, phosphorylated tau, and mitochondrial free radicals may reduce synaptic loss and enhance cognitive function in AD patients. These findings indicate the role of lysosomal function on neuronal synaptic function in the occurrence and development of AD.

5 Conclusion

In this study, two highly heterogeneous lysosomal subtypes were found in AD with significantly different immune status, autophagy status and biological functions. Lysosomes can be used as intervention targets for AD treatment in the future. This provides a further theoretical basis for the treatment of AD.

Declarations

Ethical statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was not
involved in the experiments of humans or animals.

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Declaration of competing interest

The authors have no conflicts of interest to declare.

Authors' contributions

Li Chenghao: Substantial contributions to the conception or design of the work, the acquisition, analysis, or interpretation of data for the work, and drafting the work or revising it critically for important intellectual content.


Jin Meishan: Final approval of the version to be published, agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgments

Not applicable

Funding

Not applicable

References


7. Tam JH, Seah C, Pasternak SH. The Amyloid Precursor Protein is rapidly transported from the Golgi apparatus to the lysosome and where it is processed into beta-amyloid. Mol Brain 2014, 7: 54.


**Figures**
Figure 1

(A) Volcanic plots of differentially expressed genes in AD samples and control samples. The red dots indicate that the gene expression level is up-regulated (AD sample vs. Control sample), the blue dots indicate that the gene expression level is down-regulated (AD sample vs. Control sample), and the gray dots indicate that these genes are not significantly different, with thresholds of $|\log_{2}FC| \geq 0.05$ and adjusted $p < 0.05$. (B) Heatmap displaying the expressions of top 10 up-regulated genes and top 10 down-regulated genes. Red bricks indicate the more highly expressed DEGs and blue bricks indicate lower expression. (C) Venn diagram demonstrates overlapping genes of the differentially expressed genes and Lysosome genes. Blue is the differential gene, Purple is the Lysosome-related gene.
Figure 2

ROC curves of up-regulated lysosome genes. AUC=0.5 ~ 0.7, low accuracy; AUC=0.7 ~ 0.9, with a certain accuracy; AUC > 0.9, high accuracy.
Figure 3

Identification of optimal AD subtypes based on the lysosome genes with AUC > 0.7. (A) The CDF distribution under different cluster number k, the larger the value, the more robust the clustering results under the value k. (B) The Delta area diagram shows the relative change of the area under the CDF curve between k and k-1. The larger the value is, the greater the goodness of the clustering effect under the value of k is compared with that of k-1. (C) Consistency matrix heat map with k value =2. (D) PCA clustering of cluster1 and cluster2
Figure 4

Differences in immunocell abundance between lysosome subtypes. (A) Heat map of immune cell proportions. (B) Box diagram of immune cell content.

Figure 5

(A) Venn diagram showing the overlap of genes between GSE122063 and autophagy. (B) Volcano plot displaying the expression changes of genes. (C) Heat map illustrating the expression levels of genes in different clusters.
The heterogeneity of autophagy distribution between lysosome subtypes. (A) The intersection of genes expressed in AD dataset and autophagy related genes. (B) Volcanic plots of differentially expressed autophagy related genes between lysosome subtypes. (C) Heat plots of differentially expressed autophagy related genes between lysosome subtypes.

**Figure 6**

Differentially expressed genes between lysosome subtypes. (A) Volcanic plots of differentially expressed genes between lysosome subtypes. (B) Heatmap displaying the expressions of top 10 up-regulated genes and top 10 down-regulated genes between lysosome subtypes.
Figure 7

Weighted gene co-expression network analysis (WGCNA) in the two AD subtypes. (A) Sample clustering tree. The branch represents the sample and the ordinate represents the height of the hierarchical cluster. (B) The selection of soft threshold $\beta$. The horizontal axis represents the power value of the weight parameter. The vertical axis in the left figure represents the scale-free fit index (signed $R^2$). The higher the square of the correlation coefficient, the closer the network is to the scale-free distribution. The vertical axis in the right figure represents the mean value of all gene adjacency functions in the corresponding gene model. (C) Module cluster diagram. The branch represents the module, the ordinate represents the height of the hierarchical cluster, and the red line is the selected shear height (0.2). (D) Module clustering tree diagram. The upper part is the hierarchical clustering tree of genes, and the lower part is the gene module. Corresponding to the top and bottom, the genes clustered into the same branch are divided into the same module, with different colors representing different modules. (E) Heat map of the relationship between gene modules and traits. The darker the color, the higher the correlation, with red being positive and blue being negative. The numbers in the cell represent correlation and significance. The top row is correlation, the bottom row is P-value. The left side is different colored module genes, and the right color bar represents the range of correlation.
Figure 8

Functional enrichment analysis of key module genes. (A) Bubble diagram of GO functional enrichment. (B) Bubble diagram of KEGG functional enrichment. The x-coordinate is the proportion of genes in the annotation pathway, the y-coordinate is the pathway. The color is determined by the value of p.adjust, and the size is determined by the number of genes in the annotation pathway.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterialis.rar