

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL

**AN RNA-SEQ BASED NETWORK APPROACH TO ELUCIDATE
MOLECULAR MECHANISMS OF ASYMPTOMATIC
ALZHEIMER'S DISEASE**

RÜMEYSA AKSU

**A THESIS OF MASTER OF SCIENCE
DEPARTMENT OF BIOENGINEERING
BIOINFORMATICS AND SYSTEMS BIOLOGY PROGRAM**

ADVISOR: PROF. DR. TUNAHAN ÇAKIR

FEBRUARY 2025

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL

**AN RNA-SEQ BASED NETWORK APPROACH TO
ELUCIDATE MOLECULAR MECHANISMS OF
ASYMPTOMATIC ALZHEIMER'S DISEASE**

RÜMEYSA AKSU

**A THESIS OF MASTER OF SCIENCE
DEPARTMENT OF BIOENGINEERING
BIOINFORMATICS AND SYSTEMS BIOLOGY
PROGRAM**

ADVISOR: PROF. DR. TUNAHAN ÇAKIR

FEBRUARY 2025

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

ASEMPTOMATİK ALZHEİMER HASTALIĞININ
MOLEKÜLER MEKANİZMALARINI AYDINLATMAK
İÇİN RNA-SEQ TABANLI BİR AĞ YAKLAŞIMI

RÜMEYSA AKSU

YÜKSEK LİSANS TEZİ
BİYOMÜHENDİSLİK ANABİLİM DALI
BİYOİNFORMATİK VE SİSTEM BİYOLOJİSİ
PROGRAMI

DANIŞMAN: PROF. DR. TUNAHAN ÇAKIR

ŞUBAT 2025



MASTER of SCIENCE JURY APPROVAL FORM

A thesis submitted by R meysa Aksu, defended on 25/02/2025 before the jury formed with the 10/02/2025 date and 2025/10 numbered decision of the GTU Graduate Administration Board, has been accepted as a MASTER of SCIENCE thesis in the Department of Bioengineering, Bioinformatics and Systems Biology Program.

JURY

MEMBER

(THESIS ADVISOR) : Prof. Dr. Tunahan  AKIR

MEMBER

: Assoc. Prof. Dr. Fatih TARLAK

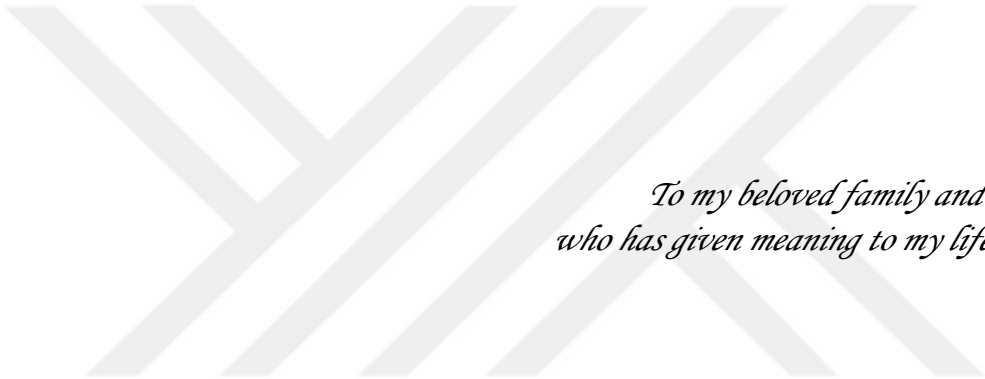
MEMBER

: Assoc. Prof. Dr. Hilal TAYMAZ N KEREL

APPROVAL

Gebze Technical University Graduate Administration Board
...../...../..... date and/..... numbered decision.

SIGNATURE/SEAL



*To my beloved family and everyone
who has given meaning to my life...*

ABSTRACT

Alzheimer's disease (AD) causes major challenges for both patients and their families due to its extremely challenging management and treatment. AsymAD (Asymptomatic Alzheimer's Disease) is a preclinical stage of AD identified by amyloid plaques and neurofibrillary tangles in cognitively normal individuals and offers essential understanding for early diagnosis and treatment of AD. The goal of this thesis research is to uncover molecular insights into AsymAD, which will lead to the identification of new candidate biomarkers and molecular mechanisms for early detection and intervention of AD. Highlighting the focus on gene expression analysis, RNA sequencing (RNA-seq) datasets from two different consortiums named ROSMAP (Religious Orders Study and Memory and Aging Project) and MSBB (Mount Sinai Brain Bank) were investigated. The individuals in the datasets were grouped into AD and AsymAD based on clinical and neuropathological criteria. Applied methods here identified differentially expressed genes (DEGs), differentially expressed transcripts (DETs) and differentially used transcripts (DUTs) between AD and AsymAD samples. Subsequently, the genes from the three analyses were mapped onto a human protein-protein interaction network using the KeyPathwayMiner tool, revealing subnetworks associated with AsymAD. The results were interpreted through functional enrichment analysis and compared with the predefined lists of AD-related and learning-memory-cognition-related genes, leading to the discovery of biomarker candidate genes. The candidate genes are NRXN3, DGKB, ADAMTS2, GNG4, ENPP5, PCOLCE, COL25A1, COL26A1, MRPL1, and MRPL30. This thesis study introduces an innovative approach by including DETs and DUTs in the analyses, beyond the standard focus on DEGs, pointing out comprehensive insights into the molecular mechanisms of AsymAD. In addition, combining the results of the subnetwork analysis from DEGs, DETs, and DUTs provided a new perspective to AsymAD and resulted in the discovery of further important genes.

Keywords: Asymptomatic Alzheimer's Disease, Differentially Expressed Genes, Differentially Expressed Transcripts, Differentially Used Transcripts, Subnetwork Analysis

ÖZET

Alzheimer hastalığı (AH), zorlu tedavi süreci nedeniyle hem hastalar hem de aileleri için büyük zorluklara neden olmaktadır. AsymAD (Asemptomatik Alzheimer Hastalığı), bilişsel olarak normal bireylerde amiloid plaklar ve nörofibriler yumaklar ile tanımlanan AH'nin prelinik bir aşamasıdır ve AH'nin erken tanı ve tedavisi için yapılan araştırmalar açısından önemlidir. Bu tez çalışmasının amacı, AsymAD'ye ilişkin moleküler mekanizmaların detaylı tanımlanmasıyla, AD'nin erken tespiti ve tedavisi için yeni aday biyobelirteçlerin tahminini gerçekleştirmektir. Bu kapsamda transkriptom verilerine odaklanılarak, ROSMAP (Religious Orders Study and Memory and Aging Project) ve MSBB (Mount Sinai Brain Bank) adlı iki farklı konsorsiyum tarafından üretilen RNA dizileme (RNA-seq) verileri analiz edildi. Veri setlerindeki bireyler klinik ve nöropatolojik kriterlere göre AD ve AsymAD olarak gruplandırıldı. Ardından, AD ve AsymAD örnekleri arasında diferansiyel olarak ifade edilen genler (DEG'ler), diferansiyel olarak ifade edilen transkriptler (DET'ler) ve diferansiyel olarak kullanılan transkriptler (DUT'ler) belirlendi. Daha sonra, bu üç analizden elde edilen genler, KeyPathwayMiner aracı kullanılarak bir insan protein-protein etkileşim ağına haritalandı ve AsymAD ile ilişkili alt ağlar tespit edildi. Sonuçlar fonksiyonel zenginleştirme analizi yoluyla yorumlandı ve literatürde yer alan AD ile ilgili ve öğrenme-bellek-biliş ile ilgili genlerin listeleriyle karşılaştırıldı. Bu sayede, aday biyobelirteç genlerin tahmini gerçekleştirildi. Diferansiyel analizlerden elde edilen aday genler NRXN3, DGKB, ADAMTS2, GNG4, ENPP5, PCOLCE, COL25A1, COL26A1, MRPL1, and MRPL30 olarak bulunmuştur. Bu çalışma, sadece DEG'lere odaklanan standart yaklaşımların ötesinde, analizlere DET'leri ve DUT'ları da dahil ederek yenilikçi bir yaklaşım ortaya koymakta ve AsymAD'nin moleküler mekanizmalarına ilişkin kapsamlı bulgular sunmaktadır. Ek olarak, DEG'lerden, DET'lerden ve DUT'lardan elde edilen alt ağ analizi sonuçlarının birleştirilmesi, AsymAD'ye yeni bir bakış açısı sağlamış ve AD için yeni genlerin belirlenmesine katkıda bulunmuştur.

Anahtar Kelimeler: Asemptomatik Alzheimer Hastalığı, Diferansiyel İfade Edilen Genler, Diferansiyel İfade Edilen Transkriptler, Diferansiyel Kullanılan Transkriptler, Alt Ağ Analizi

ACKNOWLEDGEMENTS

First of all, I would like to thank and express my gratitude to my supervisor, Prof. Dr. Tunahan Çakır, for his excellent guidance, vast knowledge and continuous encouragement. During my MSc journey, I was very lucky to have the opportunity to work with someone like him, who is hardworking and very successful in his field. I would also like to thank him for teaching me and making me apply the work discipline, apart from what he taught me about the field. Another important thanks is to my second advisor, Hatice Büşra Lüleci, PhD. When I started working with her, I was grateful for her help with everything I had difficulties with, for answering all my questions without getting tired, and for her constant moral support. Likewise, the study discipline and note-taking habits I learned from her were very helpful in the development of my thesis.

This study was supported by Scientific and Technological Research Council of Türkiye (TUBITAK) under grant number 323S202. I am thankful to TUBITAK for their support.

This thesis study utilises two large-scale datasets, which are the foundation of this thesis study, obtained from the Synapse platform (<https://www.synapse.org/#>) under the accession IDs syn17024112 and syn7416949. I acknowledge the invaluable contributions of the researchers who made these datasets available and Synapse for allowing access to these valuable datasets.

Other thanks are to my friends Evren Atak and Kazım Küçük, with whom I entered the Bioinformatics and Systems Biology Program at the same time and who I first met from my department. In addition to the good times, I spent with them, the countless homework we did and the exams we studied, I am also grateful that they introduced me to other people from the department and made me a lot of friends. I would also like to thank all my friends in the Bioinformatics and Systems Biology program and the Bioengineering department, but whose names I cannot fit to write here. I am very happy to have met all of them, and I am grateful for their support and friendships. I would also like to thank the research assistants in the Bioengineering department. Their friendships, sincerity and support were very special and meaningful to me. I'll never forget the pleasant conversations we had in the Bioengineering kitchen.

I owe a debt of gratitude to my mother, Sevim Aksu, and my father, Rafet Aksu, who have supported me through difficult times and always done their best under very difficult conditions to bring me to where I am today. I would also like to thank my brother, Abdulhamit Aksu, and my sister, Fatma Betül Aksu, for always being by my side and sharing my struggles. Finally, I extend my thanks to my ADHD medication, Concerta and its developers for helping me stay focused and attentive.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	vi
ÖZET	vii
ACKNOWLEDGEMENTS	viii
TABLE OF CONTENTS	ix
LIST OF SYMBOLS AND ABBREVIATIONS	x
LIST OF FIGURES	xi
LIST OF TABLES	xii
1. INTRODUCTION	1
1.1. Alzheimer’s Disease	2
1.1.1. Diagnosis of AD Through Clinical and Pathological Findings	4
1.1.2. Transcriptome Data and Their Usage in AD Research	5
1.1.3. Differentially Expressed Genes, Differentially Expressed Transcripts, and Differentially Used Transcripts Analyses	6
1.1.4. Network Based Analyses in AD	9
1.2. Asymptomatic Alzheimer’s Disease	10
1.2.1. Transcriptomic Studies on AsymAD	11
1.2.2. Mouse Transcriptomics Studies Representing AsymAD	14
2. METHODS	20
2.1. RNA-Seq Datasets and Data Pre-Processing	20
2.2. Grouping Datasets and Determination of the AsymAD Individuals	21
2.3. Normalisation and Covariate Adjustment	23
2.4. DEG, DET, and DUT Analyses	24
2.5. Evaluation and Integration of Differential Gene and Transcript Analyses	24
2.6. Further Investigation with Mouse Datasets	25
2.7. Subnetwork Analysis	26
3. RESULTS & DISCUSSION	27
3.1. Comprehensive Comparison of DEG, DET, and DUT Analyses Reveals Overlapping and Distinct Patterns	27
3.2. Scoring Shared Genes in ROSMAP and MSBB Datasets Compared with LMC-Related and AD-Related Gene Lists Reveals Strong Candidates	31
3.3. Subnetwork and Pathway Analyses Reveals Key Relationships and Mechanisms in AsymAD	43
3.4. Integration of Two Different Mouse Datasets to Identify and Verify Candidate Genes	54
4. CONCLUSIONS	56
REFERENCES	58
BIOGRAPHY	66
PUBLICATIONS AND PRESENTATIONS FROM THE THESIS	67

LIST OF SYMBOLS AND ABBREVIATIONS

A β	: Amyloid Beta
AD	: Alzheimer's Disease
ASEs	: Alternative Splicing Events
AsymAD	: Asymptomatic Alzheimer's Disease
AMP-AD	: Accelerating Medicines Partnership - Alzheimer's Disease
BA	: Broadmann's area
CERAD	: Consortium to Establish a Registry for Alzheimer's Disease
CDR	: Clinical Dementia Rating
CSF	: Cerebrospinal Fluid
COGDX	: Cognitive Diagnosis
DAG	: Diacylglycerol
DEGs	: Differentially Expressed Genes
DETs	: Differentially Expressed Transcripts
DGKB	: Diacylglycerol Kinase Beta
DLPFC	: Dorsolateral Prefrontal Cortex
DUTs	: Differentially Used Transcripts
ECM	: Extracellular Matrix
FP	: Frontal Pole
GNG4	: G Protein Subunit Gamma 4
KPM	: KeyPathwayMiner
LMC	: Learning-Memory-Cognition
LTD	: Long-term Depression
LTP	: Long-term Potentiation
MCI	: Mild Cognitive Impairment
MSBB	: Mount Sinai Brain Bank
NAD	: nicotinamide adenine dinucleotide
NCI	: No Cognitive Impairment
NGS	: Next-Generation Sequencing
NFTs	: Neurofibrillary Tangles
NRNX3	: Neuroxin 3
PA	: Phosphatidic Acid
PFC	: Prefrontal Cortex
PLCB1	: Phospholipase C Beta 1
PPI	: Protein-Protein Interaction
RNA-seq	: RNA Sequencing
ROSMAP	: Religious Orders Study and Memory and Aging Project
RT-PCR	: Reverse Transcription Polymerase Chain Reaction
scRNAseq	: Single-Cell RNA-seq
WGCNA	: Weighted Gene Co-Expression Network Analysis

LIST OF FIGURES

	<u>Page</u>
Figure 1.1: A workflow that summarises the main steps applied in this thesis study.	2
Figure 1.2: A characterised neuron with AD pathology.	3
Figure 1.3: Schematic summary of comprehensive comparative analyses of DEG, DET, and DUT.	9
Figure 2.1: A detailed flowchart illustrating analyses, approaches and tools employed.	20
Figure 2.2: Overview of the study workflow and the analyses before differential analyses	23
Figure 3.1: Comparison of AsymAD and AD groups and differential analysis results in both datasets.	29
Figure 3.2: The dot plot showing enrichment results of the ROSMAP genes.	30
Figure 3.3: The dot plot showing enrichment results of the MSBB genes.	31
Figure 3.4: Sankey plot showing specifically selected pathways and genes related to these pathways.	34
Figure 3.5: Evaluation of 230 candidate genes for AsymAD and AD comparison.	35
Figure 3.6: Intersection of the DET and DUT genes.	36
Figure 3.7: ADAMTS2 DEG, DET, and DUT analysis results.	38
Figure 3.8: DEG, DET, and DUT analysis results for GNG4.	40
Figure 3.9: Transcript expression and usage results of PCOLCE and ENPP5 genes, which are only significant in either the DET or DUT analyses.	42
Figure 3.10: Subnetwork result of the 230 candidate genes mapped onto a human PPI network	44
Figure 3.11: The genes identified exclusively through the transcript-based analyses (DET and DUT), not from DEG analysis, and as a consensus across the datasets.	45
Figure 3.12: The interactions between collagen biosynthesis-related genes.	47
Figure 3.13: The six subnetworks constructed based on the results of the DEG, DET, and DUT analyses from the ROSMAP and MSBB datasets.	48
Figure 3.14: Subnetwork results from DEG, DET, and DUT analyses from the ROSMAP dataset.	52
Figure 3.15: Subnetwork results from DEG, DET, and DUT analyses from the MSBB dataset.	53

LIST OF TABLES

	<u>Page</u>
Table 2.1: Pathological and clinical score definitions in ROSMAP	22
Table 2.2: Pathological and clinical score definitions in MSBB	22
Table 3.1: 230 candidate genes that are shared in ROSMAP and MSBB datasets according to the determined scoring system.	32

1. INTRODUCTION

Alongside the increase in the elderly population in the world, Alzheimer's Disease (AD), which is a devastating neurodegenerative disease, creates serious inconvenience for both patients and their relatives and healthcare institutions who take care of them (Vu et al. 2022). Despite the accumulating knowledge of AD and the variety of approaches to handling it, there is still no cure for AD. Certain treatments and approaches are mainly aimed at reducing the symptoms and slowing down the progression of the disease. In addition, earlier diagnosis of AD means more effective disease management and availability of preventive approaches to the disease (Juganavar, Joshi, and Shegekar 2023). Therefore, research on Asymptomatic Alzheimer's Disease (AsymAD), which is also known as the pre-clinical stage of AD, is crucial. Individuals with AsymAD do not show any clinical symptoms, yet they have the neuropathological burden of AD (Monsell et al. 2014). Understanding why AsymAD individuals do not show symptoms and whether there is an underlying protective mechanism in their brains is the most important and motivating starting point of this thesis study. In order to reveal such mechanisms, it is deemed appropriate to compare the individuals with AD and AsymAD and find changes between these two disease states.

In the first section of this chapter, the overall characteristics of AD, AD transcriptome datasets, diagnostic methods, transcriptomics and network-based studies in AD are reviewed. Specifically, in Section 1.1.3, three distinct transcriptomic differential analyses and the usage of these three methods together are emphasised. Studies on AD with the combinations of these three methods are also reviewed in the same section. Section 1.2 summarises the characteristics of AsymAD and various studies carried out specifically for AsymAD.

Figure 1.1 demonstrates the general design and summary of the thesis study. The details of the emphasised differential analyses are shown in Figure 1.3. The ultimate goal of this study is to reveal the underlying molecular mechanisms specific to AsymAD using the three differential analyses together, in combination with various approaches such as network and enrichment analysis. These approaches, along with

transcriptomic analyses, provide more detailed and notable insights. Additionally, two different mouse microarray datasets corresponding to AD and AsymAD status in mice were used for extra validation.

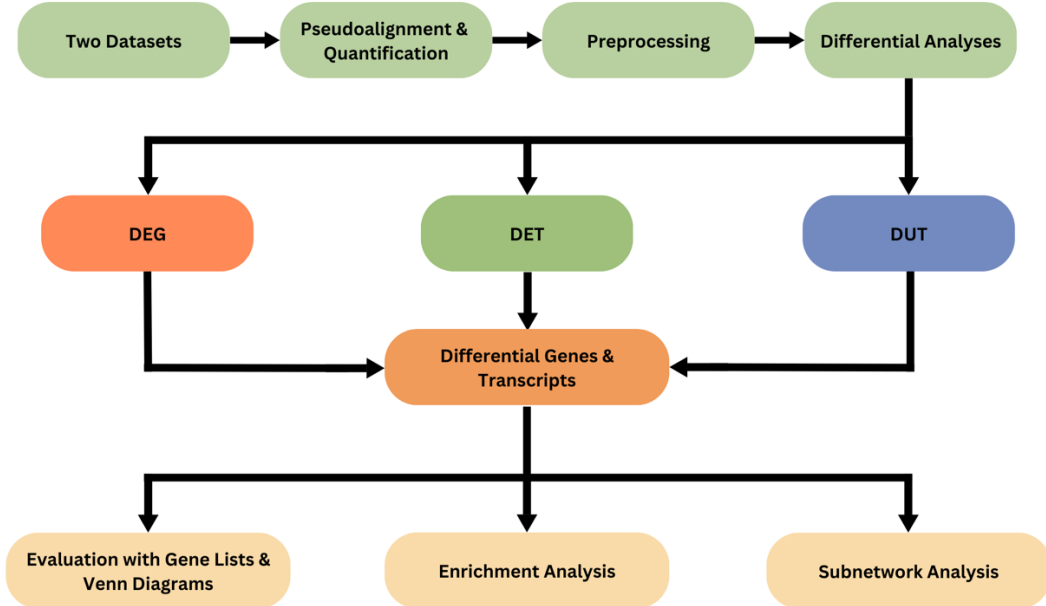


Figure 1.1: A workflow that summarises the main steps applied in this thesis study.

This study provides a detailed molecular characterisation of AsymAD, filling a large gap in the literature through the analysis of two distinct RNA-seq datasets. As a result, this study reveals new biomarker genes and pathways specific to AsymAD, which could shed significant light on understanding the molecular and pathological mechanisms underlying AsymAD and eventually lead to the potential early diagnosis of AD. The novelty of this study is the consideration of differentially expressed transcripts (DETs) and differentially used transcripts (DUTs) in the analysis, while the standard analysis in the literature is mostly based on differentially expressed genes (DEGs). At a time when AD takes a heavy toll and effective treatments are scarce, this thesis study aims to identify individuals with AsymAD who do not show obvious symptoms yet and to find efficient ways to diagnose and treat the condition.

1.1. Alzheimer's Disease

Neurodegenerative diseases often lead to gradual degeneration of the neurological system, affecting cognition, motor skills, life quality, and life span. On the other hand,

the most common neurodegenerative disease in elderly people is dementia (Ciurea et al., 2023). Dementia patients may have a significant decline in cognitive abilities until they can no longer perform their daily tasks without supervision. AD is known as the most common form of dementia. AD neuropathology is characterised by the accumulation of extracellular amyloid plaques and intracellular neurofibrillary tangles as aggregates in the brain (Figure 1.2) (Fan et al. 2020). These aggregates, which form especially in the cognitive parts of the brain, trigger a cascade of problems at the molecular and cellular levels, subsequently causing the shrinkage of the brain gradually with cognitive decline, which increases over time and eventually leads to death (Busche and Hyman 2020).

Early diagnosis of AD is crucial for early interventions and treatments that elongate the life span and delay the progression of the disease. Although some treatments can alleviate the symptoms and enhance the quality of life, it is well known that there is no decline in the symptoms once the disease starts (Knopman et al. 2021).

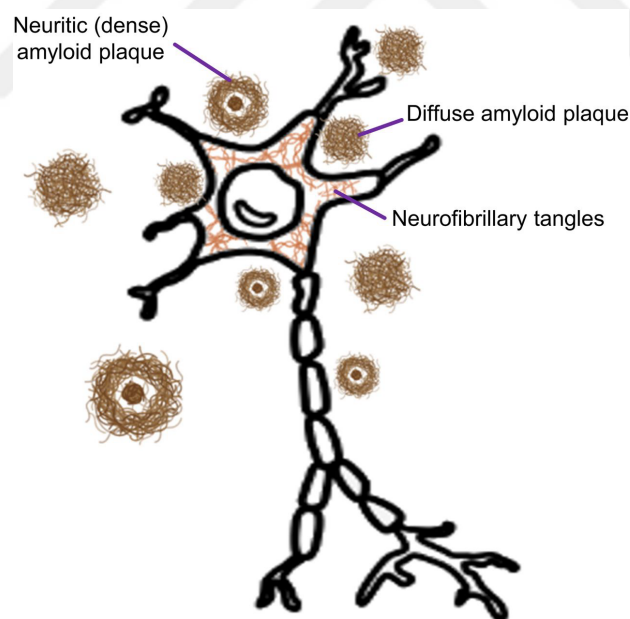


Figure 1.2: A characterised neuron with AD pathology. AD neurons include intracellular neurofibrillary tangles (NFTs) and extracellular amyloid plaques, which are either diffuse amyloid plaques or neuritic (dense) amyloid plaques that have a dense core in the centre (Gentry et al. 2024).

Diagnosis of AD might be more compelling than anticipated due to the complex nature of the disease. The complexity of AD is mostly related to the imprecision of neurological symptoms and pathologies arising from similar or comorbid diseases. Although clinical findings such as significant forgetfulness, old age, and regression in performing daily activities and mental activities are seen as the first step in diagnosing the disease, these symptoms are insufficient for the comprehensive diagnosis of AD (Robert 2013). To be precise in the diagnosis, it is quite important to check neuropathological findings in susceptible patients. Amyloid plaques and NFTs are considered to be the main pathologies. In addition, glial responses and neuronal cell loss are the other most prominent neuropathological hallmarks of AD. Amyloid plaque deposition occurs before the symptoms start, yet NFTs are more closely related to the symptoms. Microglial cells and astrocytes are mainly responsible for inflammatory responses to amyloid beta (A β) fibrils and tau tangles. Neuronal cell loss is primarily associated with synaptic loss, which is one of the primary reasons for cognitive decline (Deture and Dickson 2019).

1.1.1. Diagnosis of AD Through Clinical and Pathological Findings

To tackle the complexity of AD diagnosis, certain parameters and scoring systems based on clinical and pathological findings are used to detect and follow the progression of AD. AD diagnosis can be performed by clinicians through a multifaceted approach, starting from the assessment of cognitive and behavioural symptoms of the patients and considering their age, genetic predisposition and lifestyle, progressing to more reliable sources such as neuroimaging and in vivo biomarkers from CSF (Cerebrospinal Fluid) or PET (Positron Emission Tomography) scans (Atri 2019). Cognitive impairment is the most prominent symptom in AD hence, evaluating cognitive impairment in AD patients is the first step for clinical diagnosis. Although cognitive impairment is a strong determinant for AD, some patients may have no cognitive impairment (NCI) until death, but interestingly, the pathological burden of AD might be observed in postmortem brains.

The first example of clinical scores is the Final Consensus Cognitive Diagnosis (COGDX) score, which focuses on the cognitive performance of the patients, and they are subjected to clinical observations and some neuropsychological tests conducted. Patients are categorised into five groups according to their level of cognitive

impairment: AD, AD+, MCI (Mild Cognitive Impairment), MCI+ and NCI groups and the details of these scores are shown in Section 2.2 (Anderson 2019). MCI is a state where the transition occurs from normal ageing to dementia. Another similar clinical score is the CDR (Clinical Dementia Rating) score, which is decided by evaluating the cognitive independence of the patients based on their level of dementia. The scores are categorised into Severe to Terminal Dementia, Moderate Dementia, Mild Dementia, MCI, and NCI (Morris 1997).

Since some patients might not have any clinical signs before their death, certain neuropathological scores based on neuropathological findings compatible with AD can detect those patients in post-mortem brains. One such score is the CERAD (Consortium to Establish a Registry for Alzheimer's Disease) pathological score, and this scoring system is evaluated by considering the density of the neuritic plaques across the brain histologically (Rossetti et al. n.d.). This score is classified into four groups: Definite AD, Probable AD, Possible AD, and No AD. Another similar pathological score is Braak Staging, which falls into six stages (Stages 0-5) based upon the expansion of NFTs across brain regions (Jellinger and Jellinger n.d.). Similarly, Thal Phase scoring is based upon A β deposition across different brain regions categorised into five phases (Phases 1-5) (Thal et al. 2002).

1.1.2. Transcriptome Data and Their Usage in AD Research

Transcriptomics is the study of the transcriptome, which is the whole set of RNA transcripts in the cells or tissues for a certain developmental stage or physiological state, along with their abundance. The cutting-edge transcriptomics technology known as RNA sequencing (RNA-seq) relies on next-generation sequencing (NGS) to determine the quantity and presence of RNA in a biological sample. Other transcriptomic approaches include comparatively older technologies such as DNA/RNA microarray, reverse transcription polymerase chain reaction (RT-PCR), and Nanostring (Prokopec et al. 2013). RNA microarray is considered a pioneering technology for RNA-seq. It is a technology that allows simultaneous analysis of isolated RNAs, which are then labelled and turned into complementary DNA sequences from a sample by hybridising them into previously known DNA sequences fixed on a glass surface (chip) for a specific organism. RNA-seq is a technology considerably more favoured by scientists researching the transcriptome nowadays.

While RNA-seq is used to analyse very large sequences simultaneously, it is much more efficient in terms of both time and cost (Van Den Berge et al. 2025). Moreover, microarray can only identify previously known genes and transcripts. However, RNA-seq allows the detection of whole transcriptome by revealing novel genes or transcripts, alternative splicing events (ASEs), transcripts that show rare expressions or relatively low-abundance transcripts. Both gene and transcript-level quantification through RNA-seq can be considered one of the major advantages of RNA-seq technology (Van Den Berge et al. 2025).

Transcriptomic data plays an enormous role in AD research since it enables a better understanding of the expression patterns of the whole transcriptome to reveal underlying molecular mechanisms and pathology. In this regard, well-established publicly available RNA-seq datasets in AD are advantageous since they allow all scientists to work on AD with different objectives and approaches. The most well-recognized and large-scale datasets in AD are from the AMP-AD (Accelerating Medicines Partnership) consortium, the Religious Orders Study and Memory and Aging Project (ROSMAP), Mount Sinai Brain Bank (MSBB) and Mayo Clinic's (MayoRNAseq) datasets. Many AD studies have utilised one or more of these datasets. For instance, Mei et al. (2023) used ROSMAP, MSBB, and MayoRNAseq datasets in their study to explore mitophagy-related transcriptomic changes in the frontal lobe brain region (Mei et al. 2023). In the study of Wang et al. (2022), the researchers examined brain transcriptomic signatures linked to the progression of AD using deep learning models based on the ROSMAP, MSBB, and MayoRNAseq datasets (Wang et al. 2022). In another study, again using these three datasets, different subgroups were determined according to the transcriptomic profiles of Late-Onset Alzheimer's Disease (LOAD), and novel genetic factors affecting the pathology of AD were revealed via this classification of subgroups (Wu et al. 2023).

1.1.3. Differentially Expressed Genes, Differentially Expressed Transcripts, and Differentially Used Transcript Analyses

Differentially expressed gene (DEG) analysis is a robust and widely used technique in RNA-seq studies to identify altered gene expression levels in conditions such as disease and control individuals. DEG analysis sheds light on identifying central genes related to important mechanisms for AD pathology, including oxidative stress,

synaptic dysfunction, and immune response. For instance, some studies on extensive datasets have revealed significantly altered genes such as TREM2 in microglial activation and immune response and APOE in lipid metabolism and amyloid clearance. Furthermore, large-scale studies revealed specifically altered gene expression levels in different brain regions to emphasise regional gene expression heterogeneity of AD (Puthiyedth et al. 2016).

Transcript-level analyses such as differentially expressed transcript (DET) and differentially used transcript (DUT) analysis have been used together with classical DEG analysis in recent studies because changes in global expressions might ignore isoform-level changes, which can be important for disease mechanisms. In the human genome, each protein-coding gene has, on average, seven transcript isoforms (Sommer et al. 2022). In other words, while there are about 20,000 protein-coding genes in human DNA, the total number of defined transcripts expressed from these genes is about 140,000. DET analysis enables an in-depth understanding of underlying molecular mechanisms by considering transcript-level expression changes and identifies altered individual transcript expression levels between two conditions. In one of the recent studies, DET and DUT analyses were applied together in order to investigate transcript-level changes (Wu et al. 2023). DET analysis in that study showed the interactions between AD-specific transcript isoforms and their role in understanding AD progression by demonstrating the relationships between well-known AD pathways like synaptic function, energy metabolism and neuroinflammation. Individual transcript-level alterations of the genes obtained from the DET analysis prove transcript diversity and the involvement of these transcripts in AD pathogenesis. With the help of a transcript-level perspective, the heterogeneity of AD pathogenesis might be understood more comprehensively.

Differentially Used Transcript (DUT) analysis reveals post-transcriptional processes such as alternative splicing, alternative polyadenylation (APA), RNA editing, and alternative splice site usage by analysing the relative abundance (usage) of transcript isoforms. This approach is important for heterogeneous diseases like AD, where post-transcriptional regulations play a significant role in the pathology or progression of diseases. Farhadieh et al. (2023) used bulk and single-cell RNA-seq (scRNAseq) data to find ASEs through DUT analysis. They identified significant ASEs in neurons, astrocytes, and microglia from the prefrontal cortex (PFC) region. The study reveals

novel perspectives on cell-type-specific ASEs as they used a methodology that integrated scRNAseq and DUT analysis (Farhadieh and Ghaedi 2023). In another study, Marques-Coelho et al., 2021 used the DUT analysis approach complementarily in addition to DEG analysis to examine changes at the gene and transcript levels in AD. While DEG analysis revealed significant expression differences, especially in synaptic function, neuroinflammation, and energy metabolism genes, DUT analyses revealed transcript-specific changes that may be overlooked at the gene level, such as alternative splicing (Marques-Coelho et al. 2021). Therefore, the role of isoform-level changes and diversity in shaping disease mechanisms is critical to discover new underlying mechanisms for early diagnosis and novel therapeutic targets in AD through DET or DUT analysis.

The methods and relevant studies above show how gene-level and transcript-level changes can reveal different aspects and mechanisms behind a highly complex neurodegenerative disease like AD. A recent study applied DEG, DET and DUT analyses together to discover pathways and mechanisms in Major Depressive Disorder (MDD) at both the gene and transcript levels. They discovered sex-specific alterations in MDD patients, especially in frontal regions, through the three methods (Li et al. 2022). A representative comparison of these three methods is shown in Figure 1.3. Accordingly, some genes might be DEGs, but their transcripts might not be a DET or DUT, as demonstrated in gene A in the figure. However, individual transcriptional changes in certain genes might be lost in the traditional DEG analysis, as shown in gene B. The first graph in the figure shows that Gene A is a DEG while Gene B is not a DEG between the two conditions. The second graph shows DETs based on five transcripts, three for Gene A and two for Gene B. The first transcript, T1A, is not differentially expressed, and the remaining transcripts are differentially expressed between the two conditions according to the given values. Although Gene B is not a DEG, its transcripts are DETs. The third graph shows DUTs based on the per cent contribution of a transcript to the expression of its gene. (e.g. transcript usage of T3A is calculated as $4/8 = 0.5$ for AsymAD and $8/16 = 0.5$ for AD.) Transcript usage of T3A is not changed, while usages of other transcripts are changed between the two conditions.

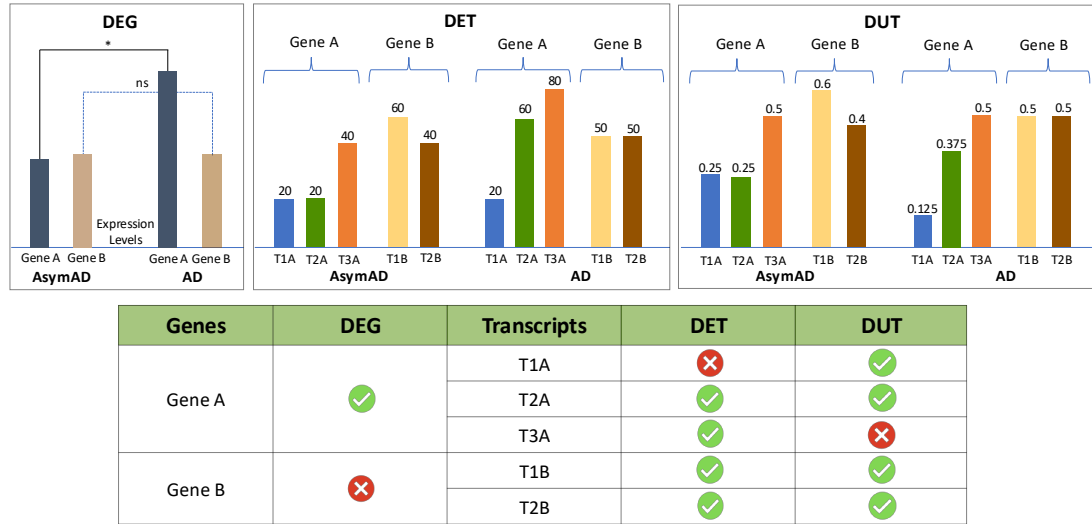


Figure 1.3: Schematic summary of comprehensive comparative analyses of DEG, DET, and DUT. * significant expression difference, ns non-significant expression difference.

1.1.4. Network-Based Analyses in AD

Network-based analyses are widely used in RNA-seq studies as they provide a comprehensive systems-level perspective on the multi-faceted nature of AD. The differential results from RNA-seq studies might not be sufficient as to whether the identified genes interact with each other in a shared molecular mechanism or pathway. For example, the interactions between resulting genes can demonstrate some functional groups according to their relevance in a specific mechanism or pathway. Gene network analysis can be performed using methods like weighted-gene co-expression network analysis (WGCNA) and protein-protein interaction (PPI) networks. Generally, after the network-based analyses, an enrichment analysis follows the workflow to gain a perspective on interaction-based subgroups in the identified networks.

The study of Düz and Çakır (2024) is a good example study of a PPI network-based analysis, as they explored subnetworks disrupted by AD. They investigated how different normalisation methods used in RNA-Seq data affect the mapping of PPI networks in AD (Düz and Çakır 2024). They used the KeyPathwayMiner (KPM) algorithm to identify subnetworks. They validated their discovered subnetworks by comparing them with known AD genes from the KEGG pathway terms related to

‘Alzheimer’s Disease’. They also applied functional enrichment analysis to the identified subnetworks, and the resulting pathways included pathways of neurodegeneration, oxidative phosphorylation system in mitochondria, response to oxygen-containing compounds, and Alzheimer’s disease. They also emphasised the importance of selecting appropriate RNA-seq normalisation methods to obtain more meaningful and precise subnetworks in AD research.

In another network-based study by Nguyen et al. (2024), they investigated common proteins from eleven studies, including hippocampus regions of AD individuals, and ultimately discovered some subnetworks and enrichment results which point out four hub proteins associated with amyloid plaques and two hub proteins with NFTs. Furthermore, they investigated common miRNAs and Transcription Factors (TFs) emerging as hub proteins from these studies and concluded that three miRNAs and two TFs could take part in the hippocampus of AD patients (Nguyen, Kim, and Huong Vu 2024).

In another network-based study by Kong et al., they integrated the PPI network and gene expression data to identify subnetworks specific and common to six different brain regions in AD. Subnetworks were optimised using the heaviest-induced subgraph algorithm and simulated annealing to construct functional modules for six different brain regions. Common subnetworks revealed that brain inflammation and calcium imbalance were key mechanisms of AD, while region-specific subnetworks revealed different biological processes. The results provided a better understanding of the role and contribution of different brain regions in AD with a network-based approach and showed that different mechanisms can operate through different brain regions (Kong et al. 2015).

1.2. Asymptomatic Alzheimer’s Disease

Asymptomatic Alzheimer’s Disease (AsymAD) is a disorder that is neuropathologically compatible with AD, yet there is no obvious cognitive decline before death (Monsell et al. 2014). Accordingly, research on AsymAD is crucial and promising for the early detection of patients who are at risk for AD and early intervention in the disease process (Driscoll and Troncoso 2011). AsymAD offers light on the multifaceted nature of Alzheimer’s disease progression. Diagnosis of AsymAD

is challenging since it is difficult to differentiate AsymAD from the normal ageing process due to the absence of cognitive symptoms. As stated in Section 1.1.1, diagnosis of AsymAD patients after death is possible by using their clinical scores (COGDX or CDR) to evaluate cognitive impairment and post-mortem pathological score based on CERAD evaluating neuropathological evidence.

Nakamura et al. (2017) investigated the early molecular alterations linked with AsymAD. This study analysed resting-state functional connectivity in 32 amyloid-negative and 13 amyloid-positive cognitively intact older adults classified by a PET imaging technique used to detect amyloid deposition. The results of the study show that functional network changes occur in the preclinical stage of AD (AsymAD) and suggest that these changes may be associated with A β accumulation. Altogether, this study advances our understanding of the preclinical stages of Alzheimer's disease by identifying specific functional network changes that occur in asymptomatic individuals at risk for Alzheimer's disease (Nakamura et al. 2017).

A study by Swarup et al. (2020) performed AsymAD to AD cell type comparisons. They found that the abundance of astrocytes and microglia increased in AD patients in relation to neurofibrillary tangle burden, yet this increase was associated with biochemical changes in AD rather than cell-type specific changes (Swarup et al., 2020). Moreover, they investigated protein networks between AD, AsymAD, and control patients with a novel proteogenomic approach. They identified differentially altered RNA-binding proteins and alternatively spliced proteins in AD and AsymAD comparison. Those alterations are more prominent in AD than in AsymAD brains, implying that RNA-related processes have a role in cognitive decline and AD pathogenesis. Also, alternatively spliced proteins are heavily related to AD risk factor proteins. Although proteomics-based studies are quite reliable, transcriptomics can detect alternative splicing events better.

1.2.1. Transcriptomic Studies on AsymAD

AD and AsymAD are sophisticated diseases that need multi-layered approaches to find out underlying mechanisms. Conventional diagnostic and therapeutic methods are insufficient to address heterogeneity and preclinical stages of AD. Multiomics approaches such as transcriptomics, proteomics, and epigenomics provide a more

comprehensive perspective on the molecular mechanisms underlying complex diseases (Hampel et al. 2021). With the help of omic studies, novel biomarkers, pathways and therapeutic targets for AD and AsymAD or mechanisms related to resilience in AsymAD can be detected. Moreover, using systems biology and network approaches enhances understanding of AD or AsymAD.

Seyfried et al. (2017) examined biomarkers and molecular alterations in AsymAD and attempted to find overlaps in their study by comparing transcriptome data and proteome data. That study used a multi-network approach to uncover protein-specific co-expression in asymptomatic and symptomatic AD. The researchers analysed brain tissue from individuals with AsymAD, AD, and controls using the WGCNA algorithm, resulting in a protein co-expression network consisting of 16 interconnected modules, and discovered that the protein co-expression network was powerful for detecting protein-specific co-expression in asymptomatic and symptomatic AD. They also used an AD microarray dataset and identified RNA-directed co-expression modules, and they compared those modules with protein co-expression networks. This study also found many protein/gene modules strongly linked to AD, including synaptic function, inflammation, and oxidative stress (Seyfried et al. 2017). However, this study did not reveal a detailed discussion on prominent pathways in AD and AsymAD comparison.

The studies above are mainly based on proteomics, but the vast majority of AD studies are transcriptomic-based due to the availability of transcriptomic datasets. Patel et al. (2019) proved the clinical importance of AsymAD based on transcriptome data. The study aimed to find DEGs in the brains of asymptomatic and symptomatic AD patients, as well as to look into possible mechanisms enabling dementia resistance in non-dementia patients with AD neuropathology. Using the microarray technology, the researchers examined postmortem brain tissues from 33 individuals with AsymAD, 52 patients with AD, and 27 age-matched controls. Together with the network analysis, they found numerous potential pathways enabling dementia resistance in non-dementia individuals with AD neuropathology, including the up-regulation of synaptic plasticity genes and down-regulation of inflammatory genes. In other words, the study provides an understanding of the molecular pathways behind AD and dementia resistance in AsymAD patients (Patel et al., 2019).

Most genes encode transcript isoforms resulting from alternative splicing, alternative transcription start site use, and post-transcriptional regulation processes such as alternative cleavage and polyadenylation. Therefore, it is critical to distinguish isoforms because they might encode proteins with distinct functions in subcellular localisations and may generate no protein product (Dick et al. 2020). Furthermore, in a study by Wu et al. (2023), transcriptional changes that occur during the progression of AD were investigated by DET and DUT analyses. The findings indicate that isoform changes and alternative splicing events play an important role in the molecular mechanisms of AD and AsymAD. In the DET analysis, they compared AsymAD and AD individuals with non-demented individuals, and they compared the DETs obtained as a result of these comparisons. They found that there were both shared and specific changes in the transcript levels. In the enrichment analysis performed based on the results of the DET analysis, pathways such as regulation of steroid biosynthetic process, regulation of lipid biosynthetic process, glucagon secretion, learning or memory, etc., were up-regulated, while nuclear export and positive regulation of insulin receptor signalling pathways were down-regulated for AsymAD. In the DUT analysis, they similarly compared AsymAD and AD individuals with non-demented individuals, and, then, they investigated isoform-switching events in these DUTs. In the enrichment analysis of genes showing isoform switching events in AsymAD, pathways such as regulation of dendrite development and regulation of amyloid-beta formation pathways were noticeable. Especially for studies at the transcriptomic level, this study offers a valuable model for biomarker discovery and detailed analysis of genetic mechanisms (Wu et al. 2023). These articles show that the combined use of DEG, DET or DUT methods provides a more comprehensive understanding of the molecular mechanisms of AD.

To summarise, the articles above have demonstrated the importance and potential of AsymAD with different approaches. Therefore, studying AsymAD is considered crucial in gaining a deeper understanding of the disease mechanism. However, these sources did not analyse DEG, DET, and DUT in a combined manner. Most of the transcriptomic studies on AD or AsymAD are DEG-based, and hence, they overlooked transcript-level changes. Additionally, although network-based approaches have been used in AD research, studies focusing on subnetworks specific to AsymAD remain limited. This dissertation fills these gaps by using DEG, DET, and DUT analyses

together and integrating network-based analyses, which leads to the identification of new candidate biomarkers and pathways associated with AsymAD. Findings specific to the AD-AsymAD comparison, which are lacking in the literature, highlight the potential molecular interactions underlying the resilience mechanisms observed in individuals with AsymAD.

1.2.2. Mouse Transcriptomics Studies Representing AsymAD

Although large-scale human transcriptomics studies are more reliable, mouse transcriptomic studies modelling the disease are more favourable and applicable. The AD mouse models that are successful in Morris Water Maze, which is a test used in mouse studies to test spatial learning, memory and behaviour, are considered a representation of AsymAD. An AD mouse model microarray dataset available in the Gene Expression Omnibus database with an ID of GSE147434, which has mice with a pathological burden and intact memory, represents AsymAD quite well in mice. The memory-impaired mice also represent AD in this dataset, and memory-impaired and memory-intact mice provide a suitable way for comparing AD and AsymAD (Pérez-González et al. 2020). In that study, they investigated DEGs between these two mice and found a highly expressed resilience gene, PLA2G4E, in intact memory mice. They also found that the AD mouse model treated with the AAV (Adeno-Associated Virus) vector significantly reversed the spatial memory deficits.

Another dataset that fits the AsymAD condition in mice is GSE111940. In cultured rat hippocampal neurons, they found that hydroxyurea (HU) protected against oxidative stress, excitotoxicity, and mitochondrial dysfunction, elevated ATP production by increasing oxygen consumption rates, and improved neuronal energy metabolism. Next, they used a mouse model for AD, and they treated half of them with hydroxyurea (HU) while the other half remained untreated as controls (Brose et al., 2018). After the microarray analysis, they analysed DEGs, and they found through Gene ontology (GO) and pathway enrichment analysis that the genes related to oxidative stress, mitochondrial function, synaptic function and neuronal energy metabolism were affected. Additionally, HU-treated mice showed attenuated spatial memory through the Morris Water Maze (Brose et al. 2018). This experimental model offers a suitable experimental design for AD-AsymAD comparison again, similar to the previous study.

2. METHODS

The methods and approaches used in this thesis study are presented in this section. Different techniques were employed and combined to obtain more valuable and comprehensive results. Figure 2.1 shows the summary of the analyses and methodology in a flowchart.

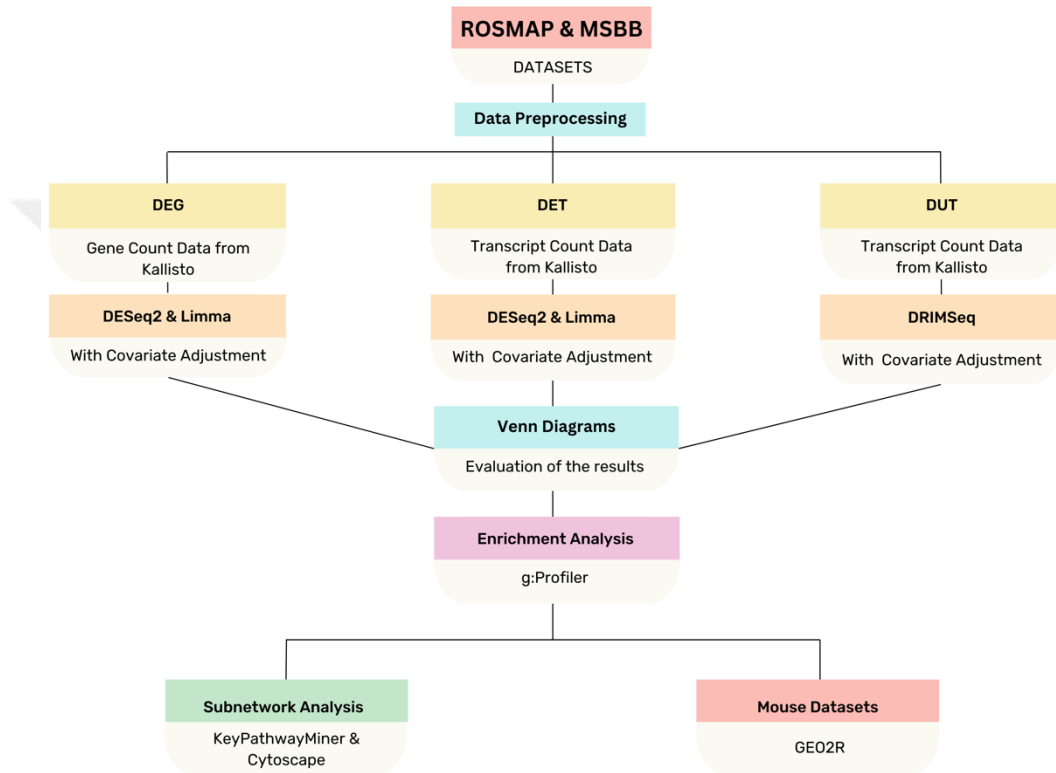


Figure 2.1: A detailed flowchart illustrating analyses, approaches and tools employed.

2.1. RNA-Seq Datasets and Data Preprocessing

The RNA-seq datasets used in this thesis study were the Dorsolateral Prefrontal Cortex (DLPFC) region of the ROSMAP (Bennett et al., 2018) and the Frontal Pole (FP) region of the MSBB (Wang et al., 2018) datasets, which are among the most commonly used RNA-seq datasets in AD. The FP region of MSBB was specifically chosen because it is located in a region where it is designated as the “cognition” region according to Broadmann’s area (BA) regions (Zilles & Amunts, 2010) in the brain (Figure 2.2.). Raw fastq reads were previously downloaded from the Synapse platform

(synapse.org) with the accession IDs syn17024112 for ROSMAP and syn7416949 and syn8612191 for MSBB, and quality-controlled and trimmed by Lüleci et al. (Lüleci et al., 2024). The pre-processed fastq files were then pseudo-aligned to the GRCh38 human reference transcriptome obtained from GENCODE (genencodegenes.org) with the Kallisto (version 0.46.1) package (Bray et al., 2016). This process involved indexing the reference transcriptome and estimating the abundance of the transcripts in the datasets using pseudo-alignment. For the subsequent DEG, DET and DUT analyses, both gene-level and transcript-level count data were created for all samples using the tximport (version 1.30.0) package.

2.2. Grouping Datasets and Determination of the AsymAD Individuals

The AD and AsymAD groups were created based on the clinical and pathological classifications in both datasets. For the ROSMAP dataset, the AsymAD group was identified using the CERAD pathological score by selecting individuals with scores 1 (Definite AD) and 2 (Probable AD) (Table 2.1, Table 2.2). Concurrently, the group was identified using the COGDX clinical score by selecting the individuals in the NCI category. Similarly, for the MSBB dataset, the pathological score was 1 and 2 according to the CERAD score, and the clinical score was defined as 0 for NCI according to the CDR score. Individuals meeting both CERAD and CDR criteria were identified as AsymAD. In both datasets, a stringent criterion for the determination of the AsymAD samples was applied by excluding the MCI samples from the clinical scores for the AsymAD group to ensure precise classification. The score table for pathological and clinical classifications is found in Table 2.1 and Table 2.2. In addition, the AD group was defined based on the consensus between clinical and pathological scoring. In other words, individuals who are categorised as AD based on both clinical and pathological scoring were used as the AD group for both datasets. As a result of these classifications, 213 AD samples and 92 AsymAD samples were determined in the ROSMAP dataset and 135 AD and 26 AsymAD samples in the MSBB dataset.

Table 2.1: Pathological and clinical score definitions in ROSMAP.

Pathological definition			Clinical Definition		
CERAD definition in ROSMAP	Scores	Definition in this study	COGDX definition in ROSMAP	Scores	Definition in this study
Definite AD	1	AD	AD+ : AD and other condition contributing to CI	5	AD
Probable AD	2		AD : AD, no other condition contributing to CI	4	
Possible AD	3	Non - AD	MCI+ : MCI and another condition contributing to CI	3	Non-AD
No AD	4		MCI : MCI, no other condition contributing to CI	2	
			NCI : No cognitive impairment	1	

Table 2.2: Pathological and clinical score definitions in MSBB.

Pathological definition			Clinical Definition		
CERAD definition in MSBB	Scores	Definition in this study	CDR definition in MSBB	Scores	Definition in this study
Definite AD	1	AD	Severe to terminal Dementia	3-5	AD
Probable AD	2		Moderate Dementia	2	
			Mild Dementia	1	
Possible AD	3	Non-AD	MCI	0.5	Non-AD
No AD	4		NCI	0	

2.3. Normalisation and Covariate Adjustment

The CPM filtration with the `cpm` function from the `edgeR` (version 3.43.8) (Robinson et al., 2009) package was used to filter out low counts and balance the library sizes. Genes with CPM values higher than 0.1 for the n number of samples, where n is the number of samples in the smaller group between the two groups to be compared, were kept in the dataset. The filtered count data were normalised with the normalisation step of the `DESeq2` (version 1.42.1) package (Love et al., 2014). Next, PCA was performed to check for any outliers in the datasets. Two outliers (Sample ID: 367_120502 and 500_120515) from the AD group were removed in the ROSMAP samples, while no outlier was removed in the MSBB samples. To reduce the effects of covariates on the data, the covariate adjustment method previously used by Lüleci et al. (Lüleci et al., 2024) was adopted. In this step, gender, age, and PMI values from the metadata were taken as covariates, and a linear model was fitted to the normalised data to identify the effects of these covariates on each gene expression value, these effects were subsequently removed. Figure 2.2 represents the brain regions in the datasets and the pre-processing steps schematically.

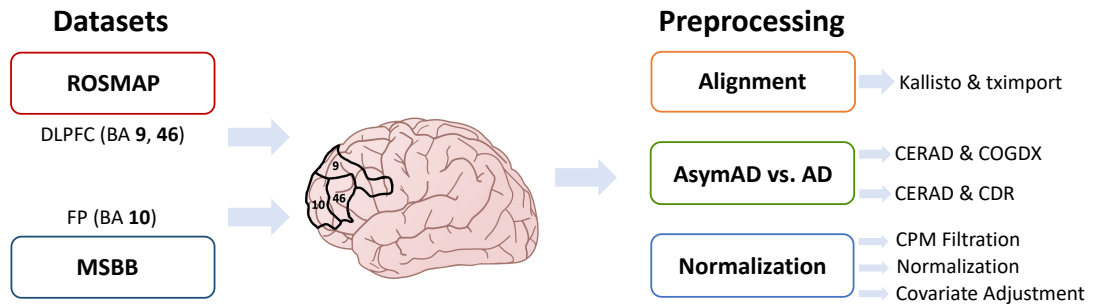


Figure 2.2: Overview of the study workflow and the analyses before differential analyses. Three different BA regions used in this thesis study derived from two datasets are shown. In the preprocessing steps, the alignment section shows the packages used to obtain gene and transcript level counts. The next section shows the grouping of AD and AsymAD samples based on their neuropathological and clinical scores. The normalisation step shows the approaches used to normalise data.

2.4. DEG, DET, and DUT Analyses

The normalised and covariate-adjusted data was then used for DEG and DET analyses with the limma (version 3.57.7) (Ritchie et al., 2015) package in R to determine differential genes and transcripts for the comparison of the two groups. The same analytical workflow was applied for both DEG and DET analyses, resulting in DEGs and DETs, respectively. A different method was followed for DUT analysis using the DRIMSeq (version 1.30.0) (Nowicka & Robinson, 2016) package in R. Additionally, the covariate adjustment method used for DEG and DET analyses was implemented in the DUT analysis by the *model.matrix* function of the DRIMSeq package. The p-value cut-off was determined as 0.001 for the ROSMAP dataset and 0.01 for the MSBB dataset for all three analyses. Subsequently, to find out which genes the identified DETs and DUTs correspond to, a biomart export file containing all human gene and transcript IDs was downloaded from Ensembl Biomart (ensembl.org/info/data/biomart/index.html), and the genes corresponding to DETs and DUTs were identified. In order to interpret the gene lists obtained from the three analyses based on the analysis of ROSMAP and MSBB datasets by identifying the relevant pathways, the gene lists were combined separately for each data set. Then, functional enrichment analysis was performed by using these two combined lists via g:Profiler (biit.cs.ut.ee/gprofiler/gost). From these two enrichment results, noteworthy pathways were selected and shown in dot plots.

2.5. Evaluation and Integration of Differential Gene and Transcript Analysis Results

The transcripts obtained as a result of DET and DUT analyses were evaluated at the gene level. Subsequently, all genes obtained from DEG, DET and DUT analyses were combined separately for the two datasets to create two different gene lists for each dataset. In these combined gene lists, the origin of the genes (i.e., from which analysis they were derived) was represented with a scoring system. According to this scoring system, genes identified in only one of the three analyses (DEG, DET, DUT) were scored as '1', those identified in two different analyses were scored as '2', and those identified in each of the three analyses were scored as '3'. Then, these scores for ROSMAP and MSBB were combined, such as R1-M2, R2-M3 etc. In other words, R1-M2 designated the case where the gene was identified in one analyses in ROSMAP,

and in two analyses in MSBB. Based on these scored lists, genes in both datasets were compared. As a result, a final gene list was created by combining the genes that were identified across the two datasets. The resulting genes were then compared with two different gene lists created previously. The first list is the AD-related gene list created by Hu et al. (Hu et al., 2017). They screened 823 genetic association studies from PubMed, and they obtained a total of 430 AD-related human genes. Those genes were then confirmed by a systems biology approach, including functional pathway enrichment analysis, pathway crosstalk analysis and subnetwork analysis. The second list, the Learning-Memory-Cognition (LMC) related gene list, was created in this thesis study from different sources by collecting genes from various articles (J. Fan et al., 2019) (Savitz et al., 2006) (Bajaffer et al., 2021) (Aimone et al., 2014) (Seto et al., 2021) (<https://dnalc.cshl.edu/>), or identifying genes related to learning, memory, or cognition from g:Profiler, and selecting human genes from the Genes to Cognition Database (genes2cognition.org). With the combination of these sources, the list includes 1542 LMC-related genes. These two gene lists were separately compared with the final scored gene list and evaluated for their significance in AsymAD. The candidate genes were then analysed with g:Profiler for enrichment analysis to reveal the functional pathways without using any enrichment criteria. Therefore, the p-value threshold in the advanced options section in g: Profiler was set to 1. From the g:Profiler result file, key pathways considered important in AD and AsymAD studies were filtered, and a separate file including those selected pathways and their corresponding genes was created. With these selected pathways and corresponding genes, a Sankey plot was created using the SRplot web tool to illustrate the interactions between the pathways and genes. (Tang et al., 2023).

2.6. Further Investigation with Mouse Datasets

For the genes that did not overlap with the LMC-related and AD-related gene lists, further analysis was conducted with two independent mouse microarray datasets, GSE111940 (Brose et al., 2018) and GSE147434 (Pérez-González et al., 2020). The DEGs for each dataset were identified for the p-value cut-off of 0.05 with the GEO2R tool (Barrett et al., 2013). In addition, these mouse datasets were compared with the DEGs obtained from the ROSMAP and MSBB datasets.

2.7. Subnetwork Analysis

Firstly, a subnetwork analysis was performed to inspect the interactions between the scored genes that were common in both datasets (See Section 2.5), and a subnetwork containing the associations of these genes was created. Secondly, apart from this subnetwork, a total of six different subnetworks were created based on DEG, DET and DTU results from both datasets. Then, the AD-related and LMC-related genes within these subnetworks were filtered into small groups within each subnetwork. Finally, for each dataset, two separate subnetworks were created by combining the AD-related and LMC-related genes identified from the DEG, DET, and DTU subnetworks: one subnetwork for ROSMAP and one subnetwork for MSBB. Thus, these two subnetworks allow visualising the general network structure of AD-related and LMC-related genes from the three analyses in both datasets separately. The KeyPathwayMiner (version 5.0) package (Alcaraz et al., 2016) was used for all of the subnetwork analyses. A PPI network containing protein-protein interactions for all species was obtained from BioGRID (BioGRID Release 4.4.236). Only interactions between human proteins were selected, and self-loops and duplicate edges within the PPI network were cleaned. The applied parameters for the KeyPathwayMiner-based subnetwork analysis were INES for the strategy, GREEDY for the algorithm, and 8 for the K parameter. The K parameter shows the allowed maximum number of non-differential genes in the subnetwork to be discovered. Cytoscape was used to visualise the networks and interpret the results.

3. RESULTS & DISCUSSION

In this thesis study, AsymAD individuals in the RNA-Seq datasets from two well-known AD cohorts, ROSMAP and MSBB, were identified based on clinical and neuropathological classifications. Then, the results of DEG, DET, and DUT analyses based on comparing AsymAD and AD groups were integrated to identify candidate biomarker genes and reveal the underlying molecular mechanisms in AsymAD. The datasets are based on the post-mortem samples from brain regions related to the prefrontal cortex of the brain. These regions were selected due to their anatomical proximity (Figure 2.2) and similar functions (e.g. learning, memory, cognition).

3.1. Comprehensive Comparison of DEG, DET, and DUT Analyses Reveals Overlapping and Distinct Patterns

AsymAD groups were not pre-determined in the datasets but were defined in this thesis study according to clinical and neuropathological scores in both of the datasets. By intersecting the individuals included in the AD group according to the neuropathological score (CERAD) and the individuals included in the NCI group according to the clinical scores (COGDX & CDR), a stringent grouping specific to AsymAD was established (Table 2.1, Table 2.2). A total of 92 AsymAD individuals for the ROSMAP dataset and 26 AsymAD individuals for the MSBB dataset were identified. In addition, the number of individuals with AD was 211 for ROSMAP and 135 for MSBB based on consensus between the neuropathological and clinical scores (Figure 3.1a).

A comprehensive comparison of DEG, DET, and DUT analyses, revealing both overlapping and different patterns between the methods, highlights the importance of the information obtained from all three methods. For instance, some of the transcripts of DEGs may be DUTs, even if they are not DETs, or vice versa. Additionally, some genes may show significant differences in transcript-level analyses, although they are not DEGs. These cases and the other combinations from these three analyses emphasise the significance of evaluating the contribution of each method, as well as the importance of evaluating the results through the integration of all three analyses.

The differential gene and transcript analyses were evaluated separately for both datasets. The results in gene and transcript levels obtained are shown in Figure 3.1b. For DEG analysis, there are 995 DEGs in ROSMAP and 320 DEGs in MSBB. For DET analysis, there are 1398 DETs in ROSMAP and 2450 DETs in MSBB, while they correspond to 992 genes (DET-G) and 1979 genes, respectively. For DUT analysis, there are 293 DUTs in ROSMAP and 541 DUTs in MSBB, and they correspond to 185 and 425 genes (DUT-G), respectively. For each analysis type, the number of common genes identified across the two datasets was 25 for DEG, 128 for DET, and 13 for DUT analyses (Figure 3.1c). One major reason behind the low overlap between the two datasets is the difference in the ROSMAP and MSBB brain regions. Despite the proximity of the brain regions. It is known that more than 95 % of human genes undergo alternative splicing, which differs depending on the type of tissue, stage of development or external signals (WANG et al. 2015). This may prove that the low similarity in the DUT analysis may be due to the region-specific ASEs. Similarly, another study examined the distribution of splicing quantitative trait loci (sQTL) in different regions of the human brain, revealing that some sQTLs are common in all brain regions while others are region-specific (Y. Zhang et al. 2020). In another study, gene expression levels were found to vary across different brain regions in AD, including BA 10 and BA 46 regions, using DEG analysis (Wang et al., 2022). One other aspect of the results reported in Figure 3 is the fact that DET and DUT analyses can capture genes that could not be identified by the commonly used DEG analysis.

To observe a clearer similarity of the results across the datasets, for each dataset, the genes from these three analyses were combined into a single list, and the similarity of the two datasets was identified by comparing these gene lists. The ROSMAP dataset results were found to have a 15% overall overlap with the MSBB dataset.

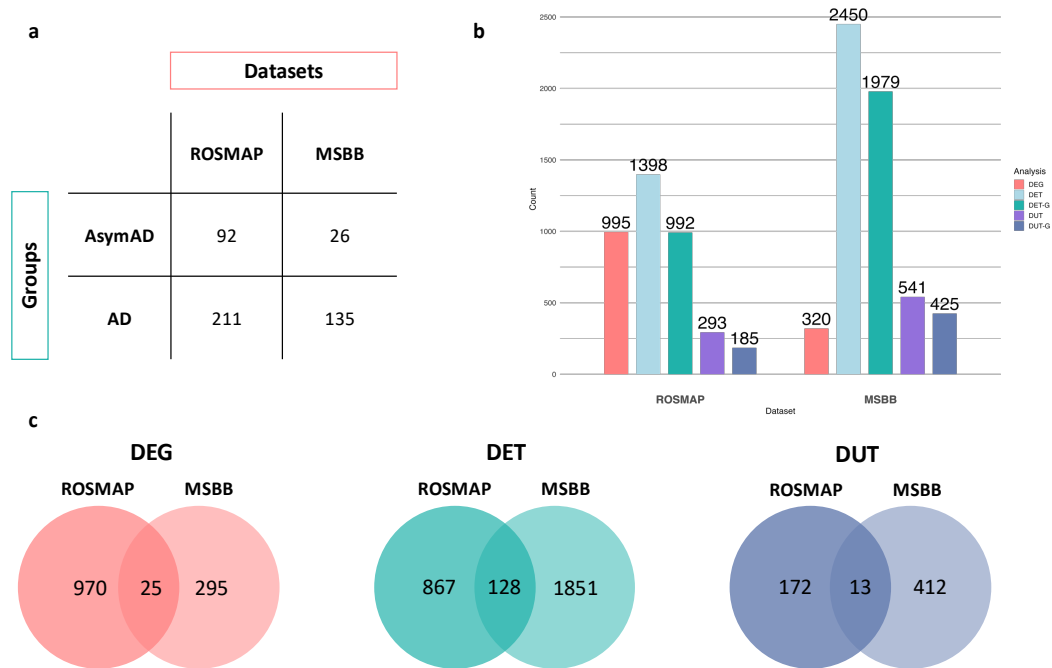


Figure 3.1: Comparison of AsymAD and AD groups and differential analysis results in both datasets. a) The ROSMAP dataset has 92 AsymAD and 211 AD individuals, and the MSBB dataset has 26 AsymAD and 135 AD individuals b) Comparison of three analysis results across ROSMAP and MSBB. c) Venn diagrams show the number of common and different genes between two datasets for each analysis.

To provide a general perspective for the ROSMAP and MSBB results and identify related pathways, combined gene lists from the three analyses were analysed in terms of enriched functional terms using g:Profiler for each dataset separately (Figure 3.2, Figure 3.3). Noteworthy findings from the ROSMAP dataset include synaptic processes and plasticity terms like regulation of synaptic plasticity, postsynaptic specialisation, presynaptic membrane, synapse organisation, postsynaptic density and synaptic vesicle cycle; neurotransmitter and synaptic function-related terms like neurotransmitter secretion, GABAergic and glutamatergic synapse, and transmission across chemical synapses; molecular and cellular regulation related terms like regulation of GTPase activity, glial cell differentiation; and cognition. Similarly, enrichment results generated from the gene list created by combining the results from the three analyses in MSBB include synapse-related processes like synapse, glutamatergic synapse, regulation of postsynapse organisation, vesicle-mediated

transport in synapse; energy metabolism and mitochondrial function-related terms like ATP binding, mitochondrial inner membrane and lipid metabolic process; and other terms such as microtubule cytoskeleton, cellular response to stress, and heme biosynthesis.

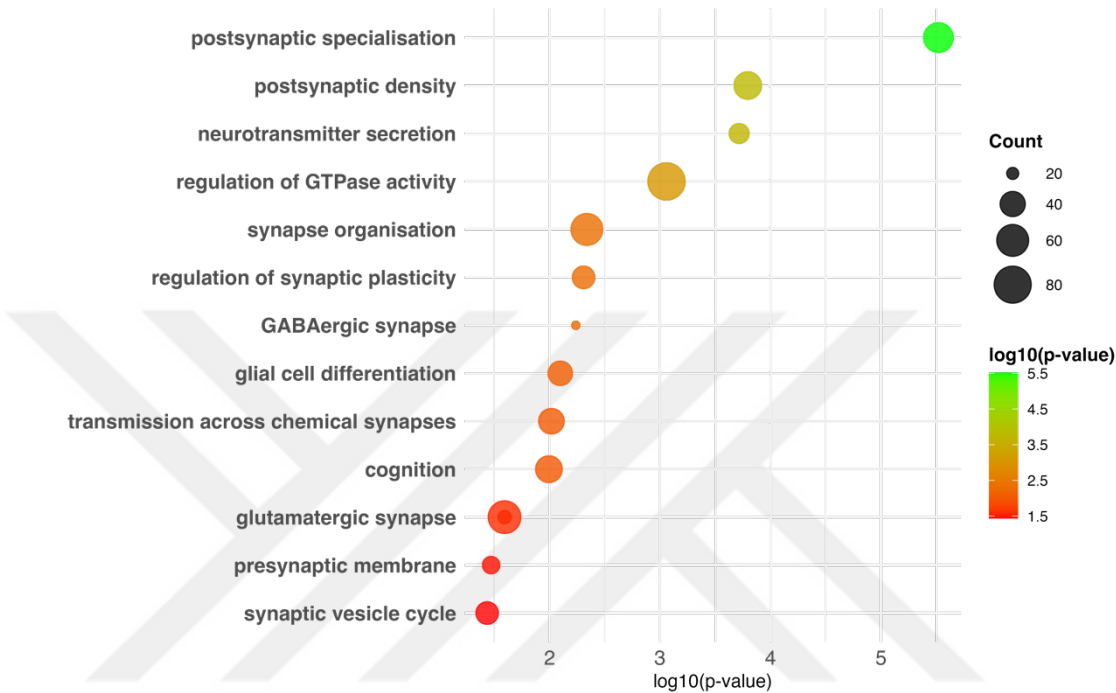


Figure 3.2: The dot plot showing enrichment results of the ROSMAP genes. The combined gene list of DEG, DET, and DUT genes was analysed via g:Profiler.

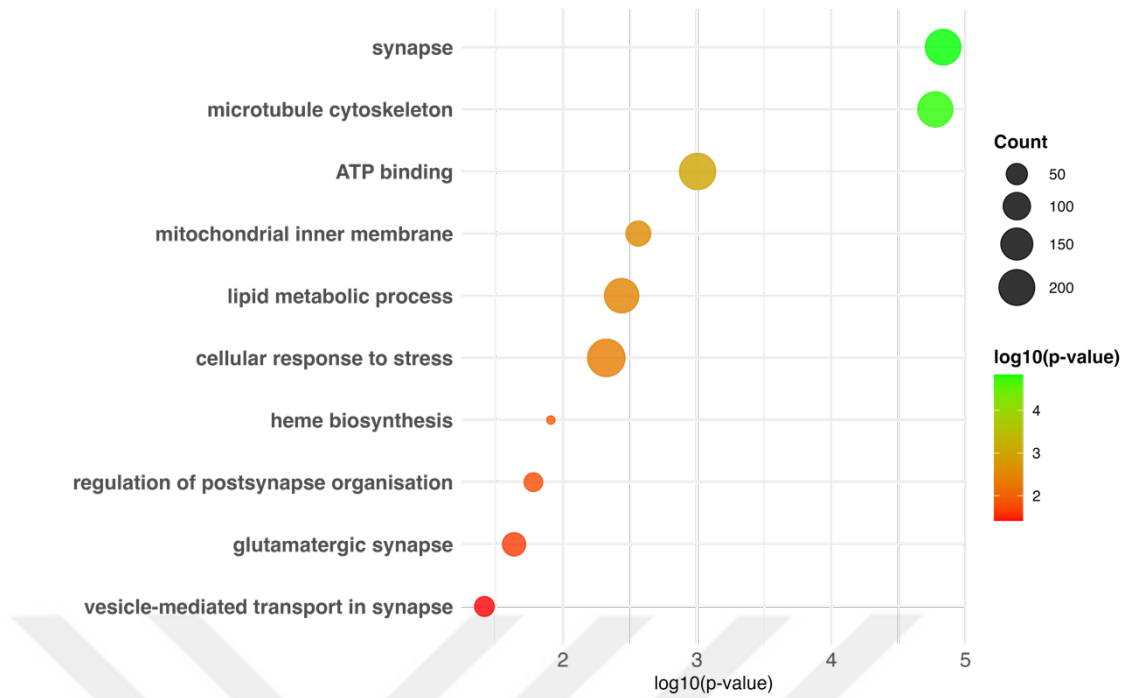


Figure 3.3: The dot plot showing enrichment results of the MSBB genes. The combined gene list of DEG, DET, and DUT genes was analysed via g:Profiler.

3.2. Scoring Shared Genes in ROSMAP and MSBB Datasets Compared with LMC-Related and AD Related Gene Lists Reveals Strong Candidates

The genes that were found to be differentially expressed or used at gene or transcript level in at least one of the three analyses in both datasets were considered to be candidate genes for AsymAD. These genes were common in both datasets, and there were 230 such genes. The scores and distributions of these 230 candidate genes are shown in Figure 3.5a. For instance, the R2-M3 group represents genes identified in at least two analyses out of the three analyses performed (DEG, DET, DUT) in ROSMAP and identified in all three analyses in MSBB. The candidate genes were distributed into nine categories based on the number of analyses that identified them in each dataset as follows: R1-M1 (153), R1-M2 (11), R1-M3 (0), R2-M1 (55), R2-M2 (8), R2-M3 (1), R3-M1 (2), R3-M2 (0), R3-M3 (0). Table 3.1 shows the 230 genes and their representative scoring groups. The bold genes in the scoring groups are highlighted due to they are present in the AD-related or LMC-related gene lists.

Table 3.1: 230 candidate genes that are shared in ROSMAP and MSBB datasets according to the determined scoring system. The genes in bold represent either AD-related or LMC-related genes.

Scores	Genes
R1-M1	AGK, GABRA1, PFKP, AP3M2, RBFOX1, PHGDH, LGMN, NAPA, WDR91, PPP3CB, FGF1, RAB3GAP1, ECE1, IQSEC2, NR1H2, CHL1, SPIRE1, DGKB, SORL1, PDXK, NDRG2, SGPL1, IDH3A, NRXN1, COL25A1, ITSN1, NME1, PDK2, TRAPPC6A, MATK, PPP5C, VEZT, HDAC9, LETMD1, SZRD1, ATP11B, POLD1, ZC3H15, WSCD2, OSBPL6, STRADB, P2RX5, MMP2, CFAP61, TASP1, FLT3LG, PSME1, RFFL, POLR2F, SOX10, PCNX1, C14orf93, NOL4, NAGPA, LMF1, TTC23, CD276, GSR, STX10, SLC1A6, PCOLCE, RPL19, CYTH1, GTF2H3, COX6A1, ENPP5, RNF130, ABCC5, MPV17, GTF3C2, UXS1, MECR, RGS4, RPS6KA1, GDA, IFI27L2, PFDN5, NDP, OVOL2, ST3GAL3, F10, PKMYT1, CDO1, ACTR3B, NIBAN2, NOVA1, RETREG3, SELENBP1, ARNT, GULP1, DENND2A, HMBOX1, FAU, DCUN1D2, THRB, MAGI1, AKAP6, RASGRP3, PTPRK, KAT6B, WDR19, CHAF1B, FTCD, PRMT2, SHKBP1, RPL29, SYNC, C1orf87, NTNG1, WDR64, FYCO1, ERMAP, DCBLD1, ZNF143, TK2, YPEL4, TERF2IP, ANKS3, DNAJC7, BMP1, NPNT, ZBBX, FAM241B, CES4A, CEP83, LRFN4, ANGEL2, VWA3A, EID2B, SNORC, SPATA13, MTA1, PDE4B, THNSL1, WWOX, MORN2, SERPINA5, CCDC9B, NBR1, MATCAP1, SLC6A9, ZNF470, H4C16, RAET1G, SYS1, LINC02210, ZNF783, ADGRG1, FXYP7, FAM229A, RPL41, ZNF37BP, GOLGA8IP
R1-M2	FBL, GCK, CDH23, UXT, CENPO, ZNF704, GNG4, PDIK1L, BRCC3, PRTN3, TCEA3
R2-M1	NRXN3, RASGRF1, RPH3A, NCALD, VAT1, SLC2A1, FGF13, GFAP, PPFIA2, CAMK4, CSRP1, KALRN, INPP5D, PLCB1, ANKS1B, YBX2, R3HDM1, HOMER3, TMEM38A, TEAD2, RBMS2, BAX, SCG3, RHOV, SLC6A12, CDK2AP1, MRPS27, PHF19, DCAF4, WDR55, PSD4, GPR108, APOL3, PNCK, ARHGEF9, RBM17, MLIP, GNA12, SERP2, TMEM178A, CARHSP1, COL26A1, NEK10, ARHGEF40, DUS2, LRRC28, WNT10B, FRMPD2, ZHX3, PDE4DIP, SPACA6, MRPL30, EPHB4, FBXO16, DMKN

Table 3.1 Continued

<i>R2-M2</i>	RIMS1, TEK2, MEIS3, PDE10A, CCDC136, PRRG3, MRPL1, CLDN5
<i>R2-M3</i>	ADAMTS2
<i>R3-M1</i>	PACSIN2 , S100A6

Some of the 230 candidate genes associated with AD or AsymAD are represented in Figure 3.4, highlighting their involvement in various pathways. Selected pathways in the Sankey plot are GTPase regulator activity, neuroinflammation and glutamatergic signalling, glucagon signalling pathway, long-term potentiation, synaptic membrane, cognition, learning or memory. The gene with the highest interaction within these pathways is PLCB1 (Phospholipase C Beta 1). This gene is linked to GTPase regulator activity, learning or memory, cognition, long-term potentiation, glucagon signalling pathway, glycerolipid metabolic process, glutamatergic synapse, cholinergic synapse, dopaminergic synapse and neurotransmitter receptors, and postsynaptic signal transmission. The highlighted gene in Figure 3.5b, NRXN3, is associated with neuroligin family protein binding, learning or memory, cognition, and the synaptic membrane. The other highlighted gene, DGKB, is less connected and is associated with the synaptic membrane and glycerolipid metabolic process. Other highly interconnected notable genes include CAMK4, PPP3CB, GNG4, RASGRF1, RPS6KA1, and NRXN1.

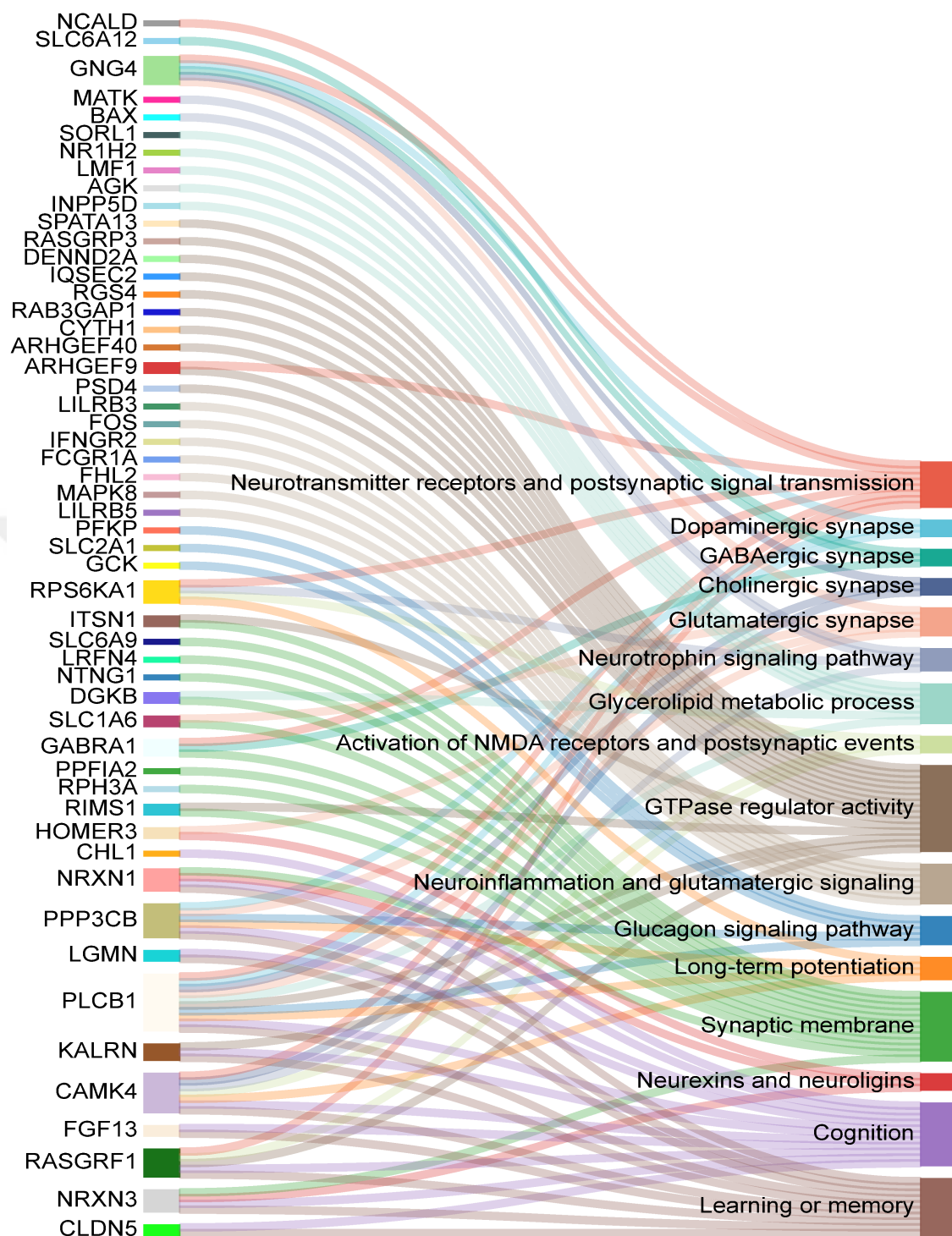


Figure 3.4: Sankey plot showing specifically selected pathways and genes related to the shown pathways.

Subsequently, these 230 candidate genes were compared with the LMC-related and AD-related gene lists. 35 LMC-related and 8 AD-related genes are common, and two genes are both LMC-related and AD-related (Figure 3.5b). These two genes are NRXN3 (Neuroxin 3) and DGKB (Diacylglycerol Kinase Beta), which are related to synaptic organisation (Bot et al. 2011) and signal transduction (Hozumi and Goto 2012), respectively, highlighting their potential role in neuronal communication and neurodegenerative processes. The NRXN3 gene, which plays a role in synaptic cell adhesion and neurotransmitter release, has previously been associated with synaptic dysfunction with lower expression in AD (Zheng et al., 2018). Similarly, the results in this thesis study showed that NRXN3 expression was lower in AD compared to AsymAD. Lower expression of this gene can interrupt synaptic integrity. The second gene is DGKB, which is related to lipid and glycerolipid metabolism. DGKB regulates diacylglycerol (DAG) signalling by converting DAG to phosphatidic acid (PA). This process is crucial for synaptic plasticity processes such as long-term potentiation (LTP) and long-term depression (LTD) mechanisms that underlie learning and memory. Loss of DGKB activity results in decreased dendritic spine density and impairs synaptic function and synaptic plasticity (Lee, Kim, and Tanaka-Yamamoto 2016). DGKB expression was identified to be lower in AD than in AsymAD in this thesis study, and this shows that higher expressions of this gene might be related to synaptic plasticity, hence neuroprotection.

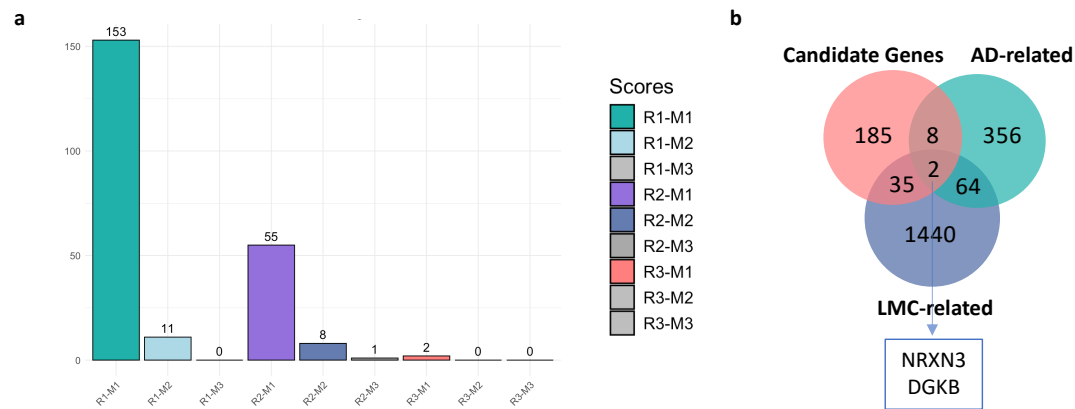


Figure 3.5: Evaluation of 230 candidate genes for AsymAD and AD comparison. a) Genes derived from DEG, DET, and DUT analyses were evaluated across nine combinations for the two datasets. b) A Venn diagram shows the comparison of 230 candidate genes with AD-related and LMC-related gene lists.

Additionally, in Figure 3.6, only the intersections of transcripts in DET and DUT analyses were compared across two datasets to reveal the differential transcripts that were compatible across the two datasets. It was found that eleven transcripts were common in the DET analysis, and four transcripts were common in the DUT analysis. Three of the transcripts of the DET analysis did not match with any gene. Possible reasons for the transcripts with no gene assignment are the availability of uncharacterised genes or novel non-coding genes.

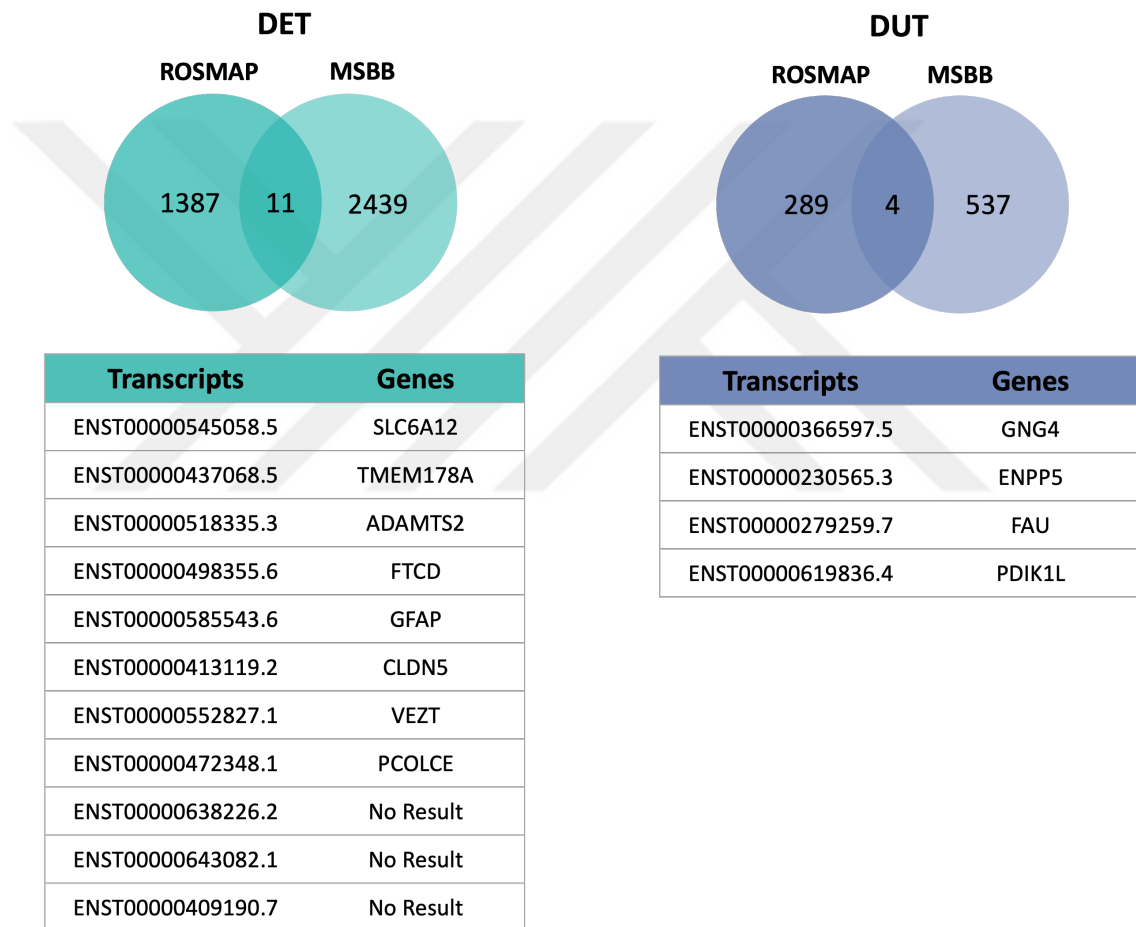


Figure 3.6: Intersection of the DET and DUT genes. Tables from the DET and DUT analyses represent the shared transcripts and their corresponding genes between the datasets.

The first example from these genes is ADAMTS2, which was identified as DEG and DET for ROSMAP and as DEG, DET, and DUT for MSBB (Figure 3.7). A disintegrin-like and metalloproteinase with thrombospondin type 1 motif (ADAMTS) is a family

of extracellular matrix (ECM) proteases (Rose et al. 2021). There is evidence that reelin, known to regulate neurogenesis, synapse formation, and synaptic plasticity positively, is degraded by ADAMTS and that decreased reelin levels may increase tau phosphorylation (Gurses et al. 2016). A study involving the generation of ADAMTS2 knockout (KO) mice found that ADAMTS2 significantly contributes to the proteolytic cleavage and inactivation of reelin in the postnatal cerebral cortex and hippocampus (Yamakage et al. 2019). In addition, another study using the ROSMAP cohort found that high expression of ADAMTS2 accelerated cognitive decline in episodic memory and global cognitive functions (McCorkindale et al. 2022). In the results of this thesis study, the global expression of the ADAMTS2 gene, classified within the R2-M3 group, increased in both datasets in AD. The gene was detected in the MSBB dataset as DEG, DET, and DUT and in the ROSMAP dataset as DEG and DET, making it a robust candidate. In the DET case, it was observed in this study that the expression of one of its isoforms, ADAMTS2-203 (ENST00000518335.3), increased in both datasets in the AD condition. In the DUT analysis case, its proportions increased within AD in ROSMAP. The lower activity of the ADAMTS2 gene, especially the ADAMTS2-203 transcript, may provide a neuroprotective effect in AsymAD. At the same time, its overexpression in AD may be directly associated with AD neuropathology.

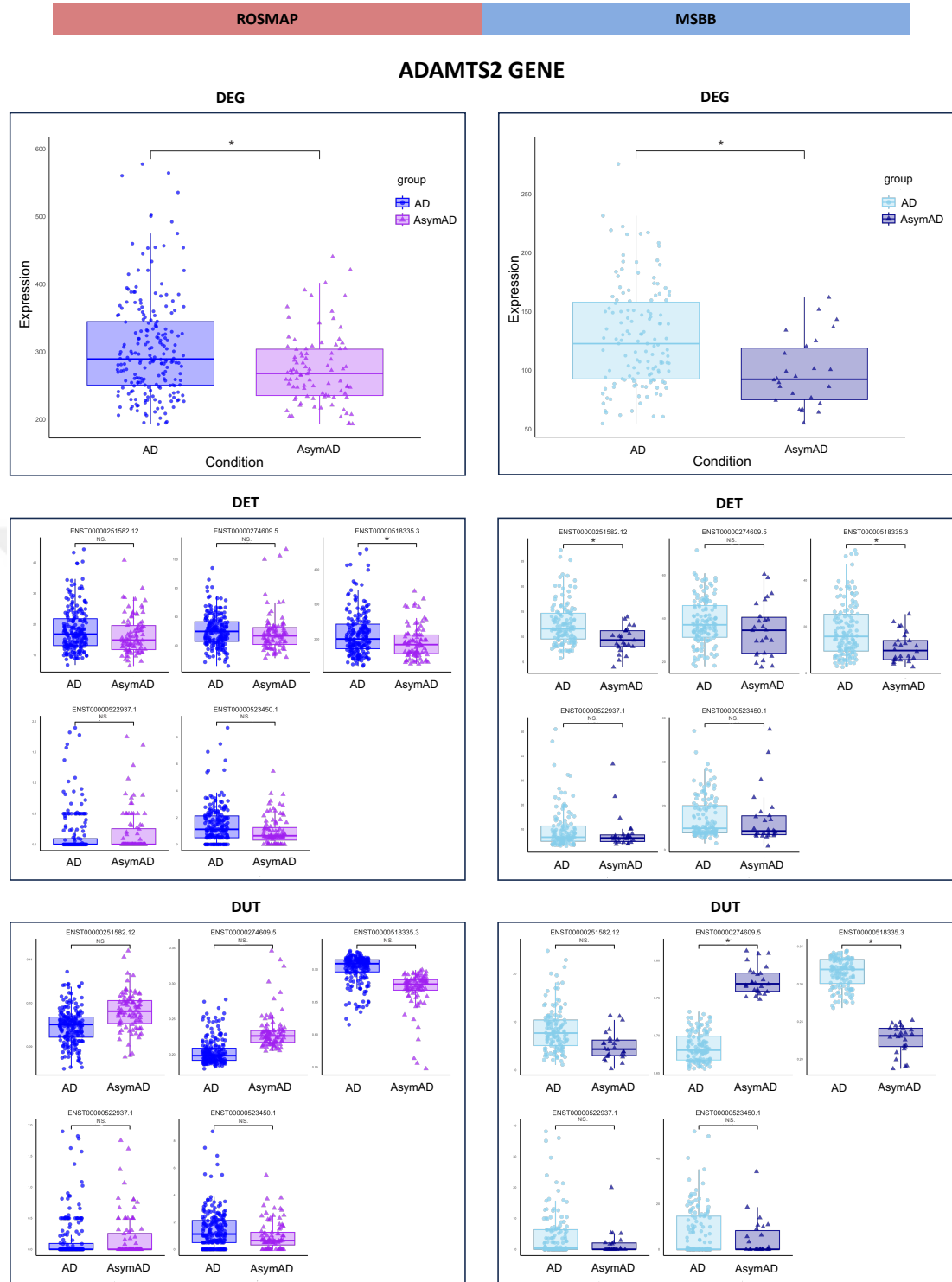


Figure 3.7: ADAMTS2 DEG, DET, and DUT analysis results. Gene expressions, transcript expressions, and transcript usages of the ADAMTS2 gene are significantly different between AD and AsymAD groups for both DET and DUT analyses in the two datasets.

The second one (Figure 3.8) is GNG4, which was identified only in the DUT analysis for ROSMAP and in the DET and DUT analyses for MSBB. GNG4 (G protein subunit gamma 4) is associated with pathways such as neurotransmitter receptors and postsynaptic signal transmission. Bonham et al. (2018) suggested that GNG4 may be related to cognitive decline during normal ageing. In that study, GNG4 showed relatively higher expression in the frontal cortex. GNG4 expression was reported to decrease with advancing age (Bonham et al. 2018). In another study, significant transcriptional changes in synaptic function-related genes were observed, and the GNG4 gene was downregulated in 12-month-old 5XFAD mice (Bouter et al. 2014). Bayraktar et al. (2021) investigated the molecular mechanisms of AD through network analysis, and extensive dysregulation of synapse-associated genes was observed in AD. The results showed that GNG4 was downregulated in DLPFC (Bayraktar et al. 2021). Interestingly, in the results obtained in this thesis study, the overall expression of GNG4 was not significantly changed in the DEG analysis. Yet, one of its transcripts, GNG4-201 (ENST00000366597.5), is up-regulated in the MSBB dataset, and the proportion of that transcript is higher in the AD condition for both datasets. The fact that this transcript is up-regulated in AD within the MSBB dataset and the transcript proportion increases in both datasets in a compatible manner suggests that the increase in the expression or proportion of this isoform of the GNG4 gene may be associated with AD pathology.

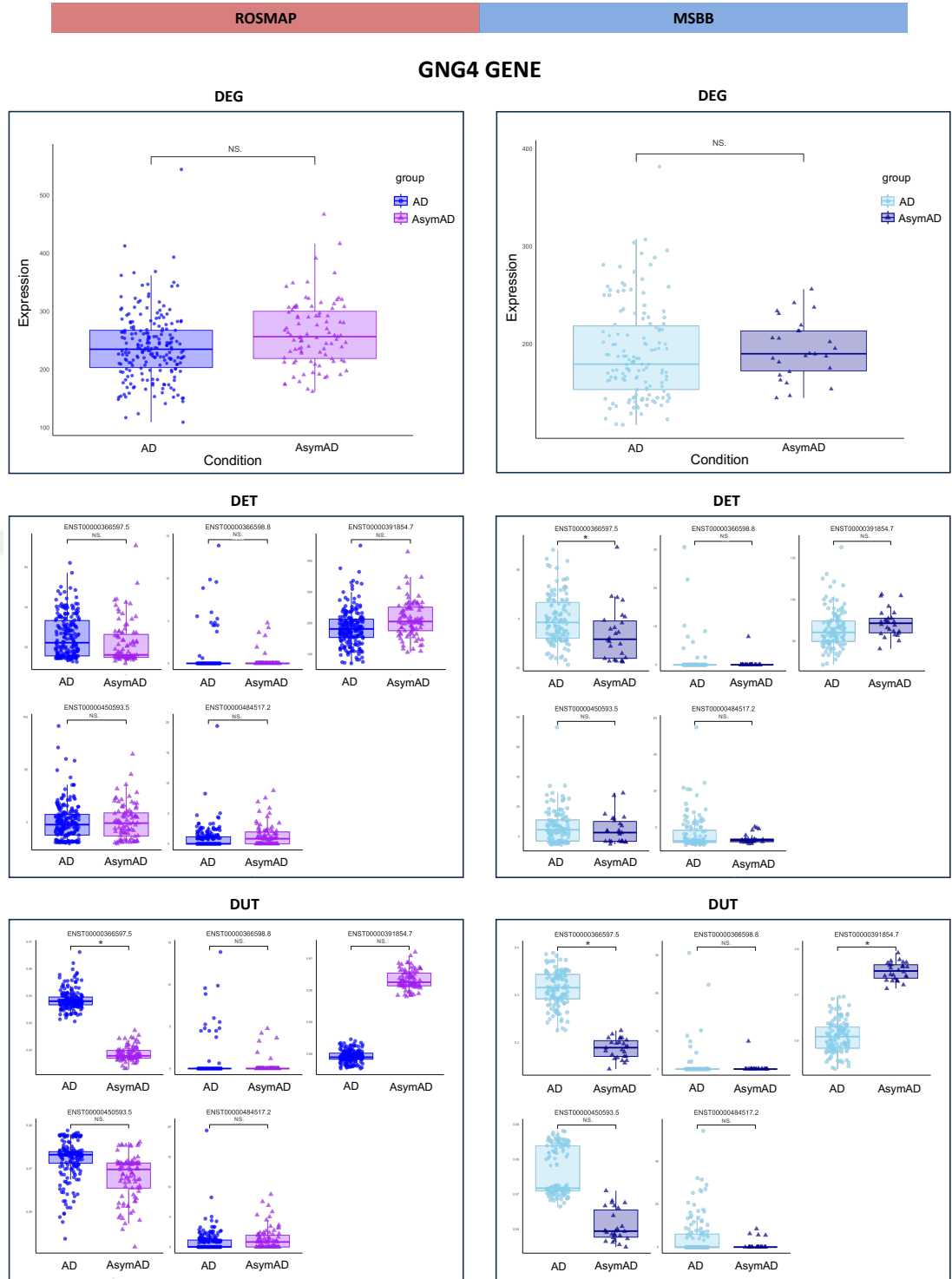


Figure 3.8: DEG, DET, and DUT analysis results for GNG4. Gene expressions, transcript expressions, and transcript usages of the GNG4 gene are significantly different between AsymAD and AD groups in both datasets.

Another gene was identified to be a common DUT between the two datasets PCOLCE (Figure 3.6). One of the transcripts of the PCOLCE gene (ENST00000472348.1) was down-regulated in AsymAD (Figure 3.9a), while in DUT, the proportion of a transcript of the ENPP5 gene (ENST00000230565.3) is higher in AsymAD (Figure 3.9b) in both datasets. The PCOLCE (procollagen C-endopeptidase enhancer 1) gene is a member of the collagen family gene located in the ECM. The study by Lepelletier et al. (2017) observed alterations in the expression of ECM molecules during the early stages of AD (Lepelletier et al. 2017). In addition, another study found that collagen-binding proteins, including PCOLCE, are dysregulated in the PFC of AD (Williams, Cao, and Yan 2021). In this thesis study, a non-protein-forming, intron-retained transcript of the PCOLCE gene, PCOLCE-205 (ENST00000472348.1), was down-regulated in AsymAD. This may prove that the reduction of this transcript, specific to AsymAD, has a protective effect. In addition, it is known that the ADAMTS2 and PCOLCE genes encode enzymes that modify procollagen molecules (Smagin et al. 2019). The fact that these two genes are collagen family genes emphasises the importance of this pathway, especially the role of these genes in AD. ENPP5 (Ectonucleotide Pyrophosphatase/Phosphodiesterase Family Member 5) is the least characterised ENPP family member. It has been suggested that it degrades nicotinamide adenine dinucleotide (NAD) and may have a potential role in NAD-based neurotransmission, and studies in rats indicated that the encoded protein may play a role in neuronal cell communication. A transcript isoform of the ENPP5 gene, ENPP5-201 (ENST00000230565.3), was found to have an increased proportion in the AsymAD group in both datasets. The increased proportion of the ENPP5-201 transcript isoform in AsymAD compared to AD may be a compensatory response to balance NAD metabolism in the early stages of the disease. Since decreased NAD levels can accelerate neurodegenerative processes (Blacher et al. 2015), an increased proportion of ENPP5-201 in AsymAD may function as a response to oxidative stress or as an early cellular defence mechanism (Gong et al. 2013). However, in the later stages of AD, a decrease in the proportion may be associated with a general degradation of NAD metabolism, a change in the proportions of different transcript isoforms, or altered synaptic functions (Chaubey et al. 2022). Thus, while the increased proportion of ENPP5-201 in AsymAD may help maintain cellular homeostasis in the early stages of the disease, it may indicate that compensatory mechanisms become inadequate due to changes in the overall transcript balance in the later stages.

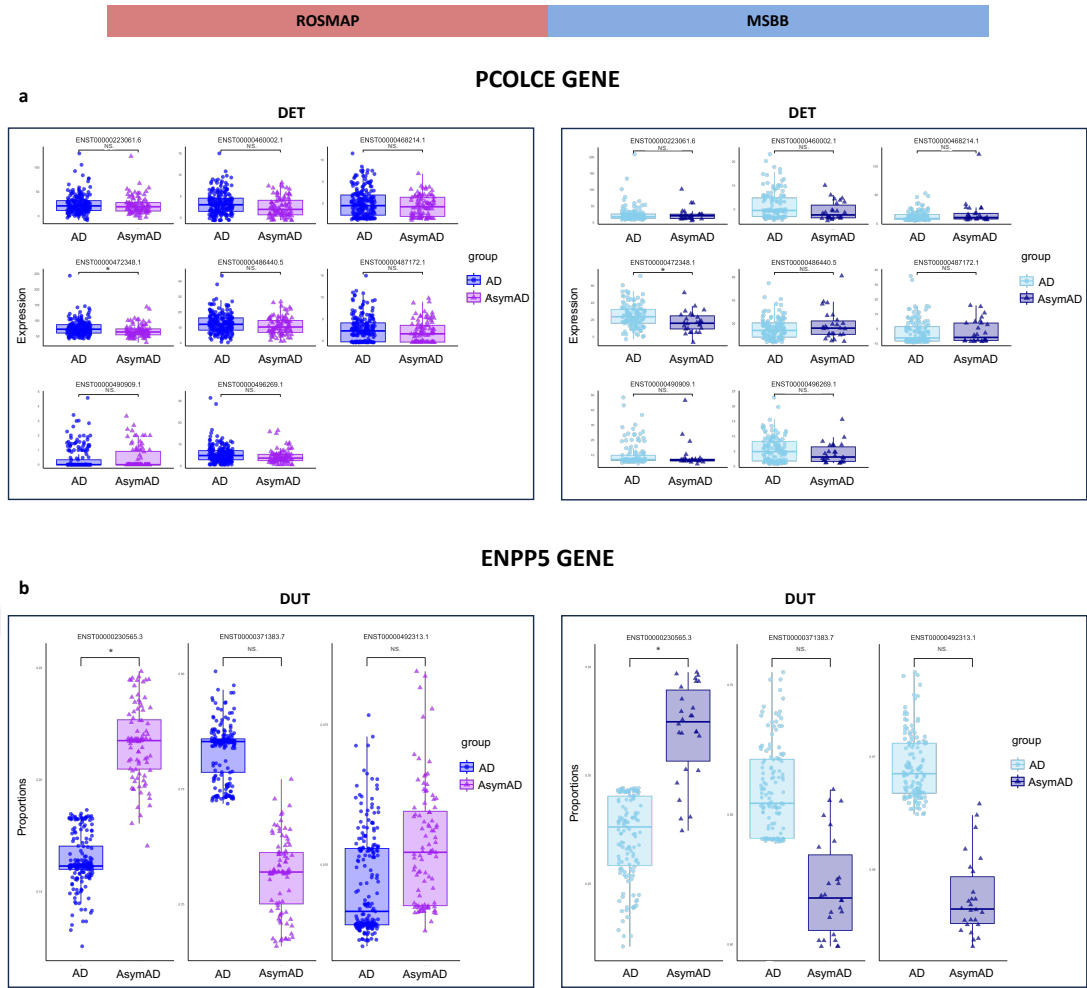


Figure 3.9: Transcript expression and usage results of PCOLCE and ENPP5 genes, which are only significant in either the DET or DUT analyses. a) Transcript expressions of the PCOLCE gene are significant for the DET analysis in both datasets. b) Transcript usages of the ENPP5 gene are significant for the DUT analysis in both datasets.

3.3. Pathway and Subnetwork Analyses Reveal Key Relationships and Mechanisms in AsymAD

In Figure 3.10, mapping the 230 candidate genes to the human PPI network revealed that 176 of these genes interacted with one another, forming a subnetwork. Genes marked in grey are not part of the original list, but their major function is to interconnect the genes in the list. APP is one of such genes, and it has very high numbers of interactions and is strongly associated with AD. In the R3-M1 group, PACSIN2 and S100A6 were identified as strong candidates, significant in three analyses in ROSMAP and one in MSBB. In the R2-M2 group, RIMS1, MRPL1, and CLDN5 were highlighted. RIMS1 was associated with synaptic membrane and GTPase regulator activity (Figure 3.4). In the R2-M1 group, CAMK4, PLCB1, COL26A1, NRXN3, RASGRF1, MRPL30, FGF13, and GFAP were highlighted. CAMK4 is associated with pathways such as learning or memory, cognition, long-term potentiation, activation of NMDA receptors and postsynaptic events, neurotrophin signalling pathway, Cholinergic synapse and neurotransmitter receptors, and postsynaptic signal transmission (Figure 3.4). In the R1-M2 group, GNG4 and FBL genes were featured in the subnetwork. GNG4 was associated with pathways including glutamatergic synapse, cholinergic synapse, GABAergic synapse, dopaminergic synapse, neurotransmitter receptors and postsynaptic signal transmission (Figure 3.4).

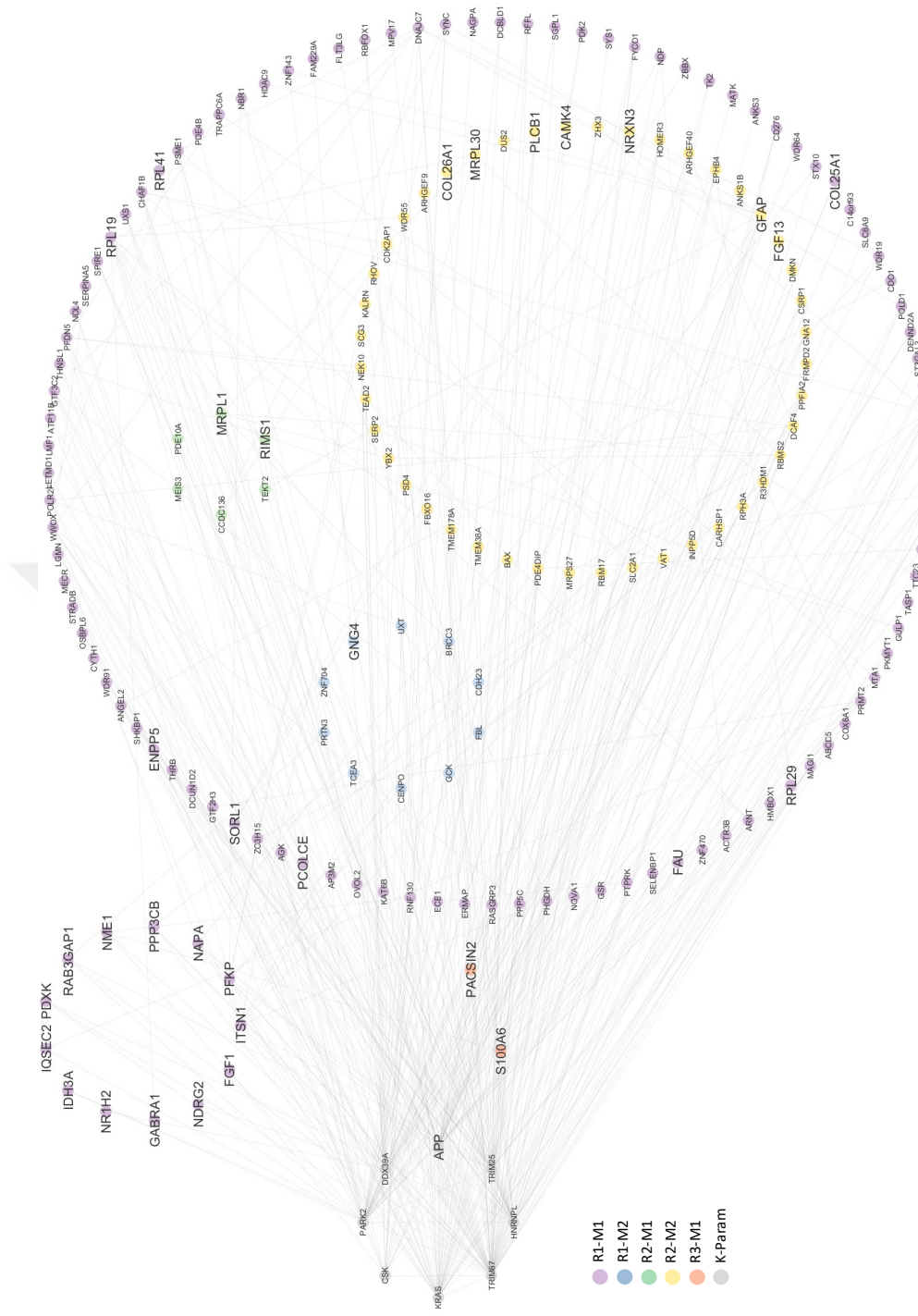


Figure 3.10: Subnetwork result of the 230 candidate genes mapped onto a human PPI network, 176 genes mapped and exhibited strong interactions with each other.

A subset within the R1-M1 group represents genes that were not identified in the DEG analyses but significant in the transcript analyses (DET and DUT) and also part of the LMC-related and AD-related gene lists. Those genes are GABRA1, NR1H2, ITSN1, NME1, PFKP, RAB3GAP1, PDXK, FGF1, NAPA, NDRG2, IQSEC2, PPP3CB, and IDH3A. Those genes were obtained by first identifying genes commonly determined as DETs in both datasets and genes commonly determined as DUTs in both datasets (Figure 3.1c). Then, genes that were in the DEGs lists of either of the datasets were excluded from the common DET and common DUT genes (Figure 3.11). These genes prove that transcriptomic studies relying solely on DEG analysis can miss transcript-level changes. The other genes highlighted in the R1-M1 group outside this category are COL25A1, PCOLCE, RPL19, and ENPP5.

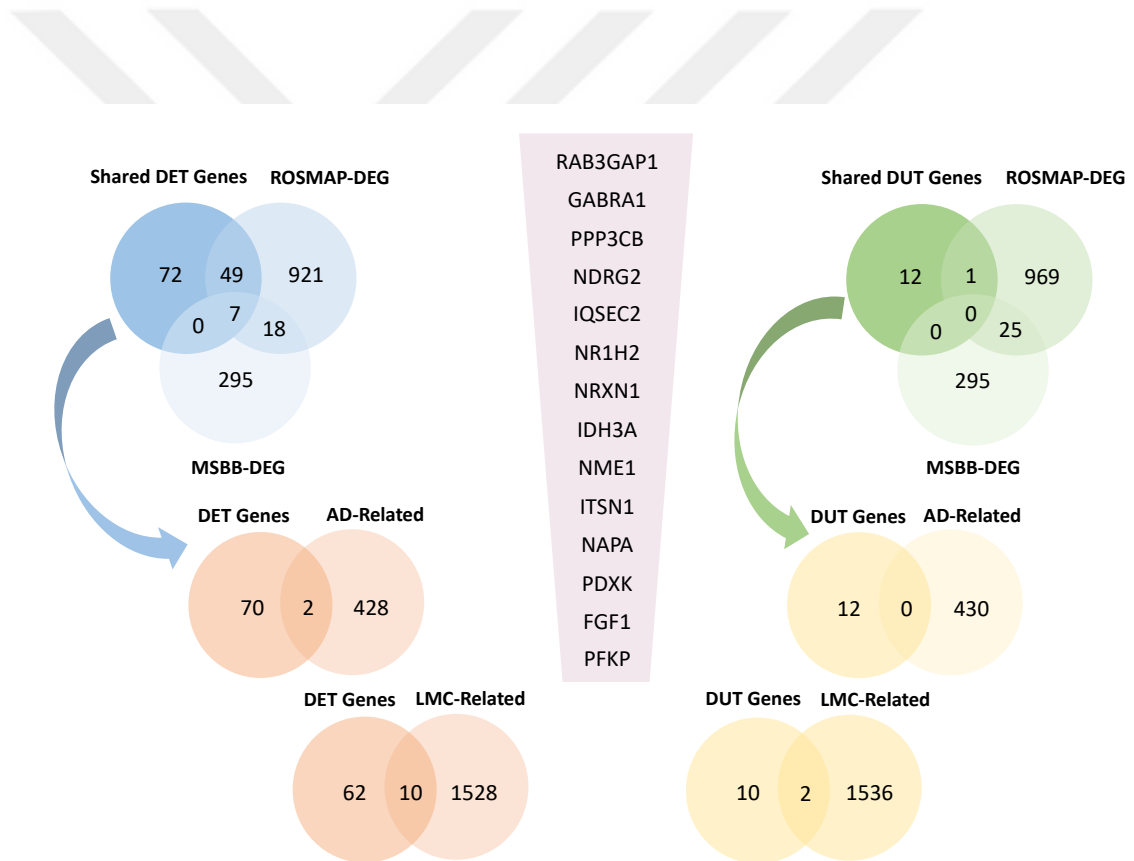


Figure 3.11: The genes identified exclusively through the transcript-based analyses (DET and DUT), not from DEG analysis, and as a consensus across the datasets. Those genes were also compared with the AD-related and LMC-related gene lists.

Some protein family genes were also identified in the discovered subnetwork. For example, the COL25A1, COL26A1, and PCOLCE genes belong to the same protein family, as they encode collagen family proteins (Smagin et al. 2019). PCOLCE may be a critical regulator that indirectly affects the functions of collagen family genes COL25A1 and COL26A1. What these genes have in common is their effects on the structuring and organisation of the ECM. ECM has essential functions in the central nervous system, such as structural support, intercellular communication, and extracellular signalling. Abnormal accumulation and modifications of ECM components in AD may contribute to the formation of $\alpha\beta$ plaques and tau proteins and neurodegeneration processes through synaptic signal transmission, oxidative stress and inflammatory responses (Sun et al. 2021). Since COL25A1 plays an important role in the neurodegenerative processes (Tong et al. 2010), COL26A1 in the connective tissue stability (Bader et al. 2009), and PCOLCE in the maturation of procollagens (Steiglitz et al. 2006), dysfunction of these genes may lead to diseases like AD, related to the disruption of the integrity of the ECM. Also, the enrichment results show that the ‘collagen biosynthesis and modifying enzymes’ term is associated with the genes ADAMTS2, COL26A1, PCOLCE, BMP1, and COL25A1. Among these genes, PCOLCE, COL25A1, and COL26A1 interacted with each other not directly but through the exception genes, the genes that appeared in the subnetwork due to the K parameter in the KPM algorithm (Figure 3.12), although they were not in the DEG, DET or DUT lists. The exception genes played a critical role in the subnetwork discovery by bridging the input genes. Pathway enrichment of the K-parameter group indicates that the genes within this group are primarily associated with terms such as cytoplasm, regulation of intracellular organelle lumen, signal transduction, response to stress, and apoptotic process as highly related terms, along with cognition, astrocyte development, and regulation of spontaneous synaptic transmission as less related terms. Another example of such related genes that are in the same protein family are MRPL1 and MRPL30. They are genes that encode proteins associated with the large subunit of the mitochondrial ribosome. These genes play critical roles in mitochondrial protein synthesis and support cellular energy production. Disruptions in the function of mitochondrial ribosomal proteins are associated with oxidative stress and dysfunction in energy metabolism and may play a role in AD (Lunnon et al. 2017). Furthermore, MRPL30 is one of the genes validated by the mouse datasets used in this thesis study (See Section 3.5). Additionally, cytosolic ribosome genes such as RPL19,

FAU, RPL29, and RPL41 can serve as similar examples. Due to their ribosomal functions and critical roles in protein synthesis, these genes may contribute to oxidative stress, energy metabolism disorders, and cellular stress mechanisms in AD. Unbalanced expression of ribosomal proteins in the brain is associated with neuronal loss and dysfunction observed in AD (M. Suzuki et al. 2022).

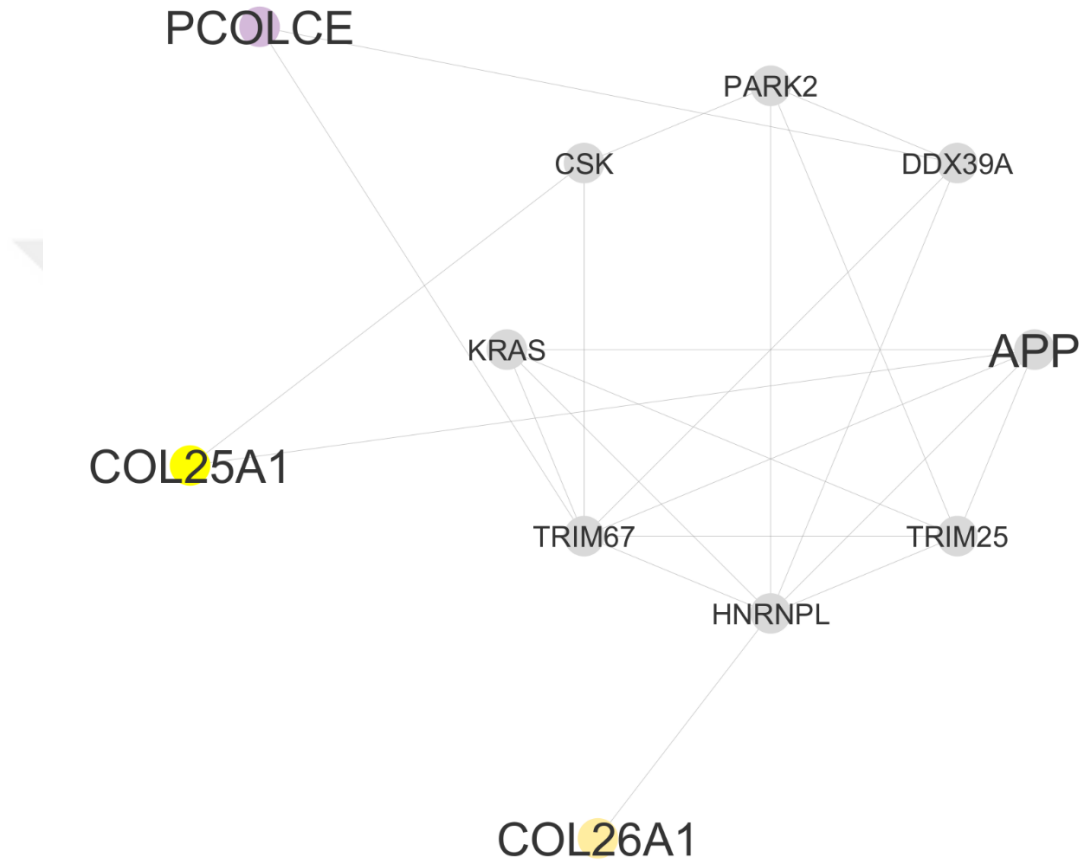


Figure 3.12: The interactions between collagen biosynthesis-related genes. The interactions between ‘collagen biosynthesis and modifying enzymes’ pathway-related genes and their interactions with the genes that are not part of the original 230 candidate genes but the genes included in the subnetwork due to the K-parameter in the KPM algorithm.

The six subnetworks representing DEG, DET, and DTU results from the ROSMAP and MSBB datasets are shown in Figure 3.13. The three individual subnetworks from the DEG, DET and DTU analyses in ROSMAP resulted in 747 nodes and 3067 edges for the DEGs, 749 nodes and 3009 edges for the DETs and 156 nodes and 307 edges

for the DUTs. Moreover, the three individual subnetworks from the DEG, DET and DTU analyses in MSBB resulted in 108 nodes and 162 edges for the DEGs, 1463 nodes and 9187 edges for the DETs and 304 nodes and 762 edges for the DUTs. The genes in the small circles illustrated on the right side of these six subnetworks were created by filtering out the AD-related and LMC-related genes. The three small circles on the right, filtered for the ROSMAP dataset, were merged into the ROSMAP subnetwork (Figure 3.14). Likewise, the three small circles on the right, which were filtered for the MSBB dataset, were merged into the MSBB subnetwork (Figure 3.15).

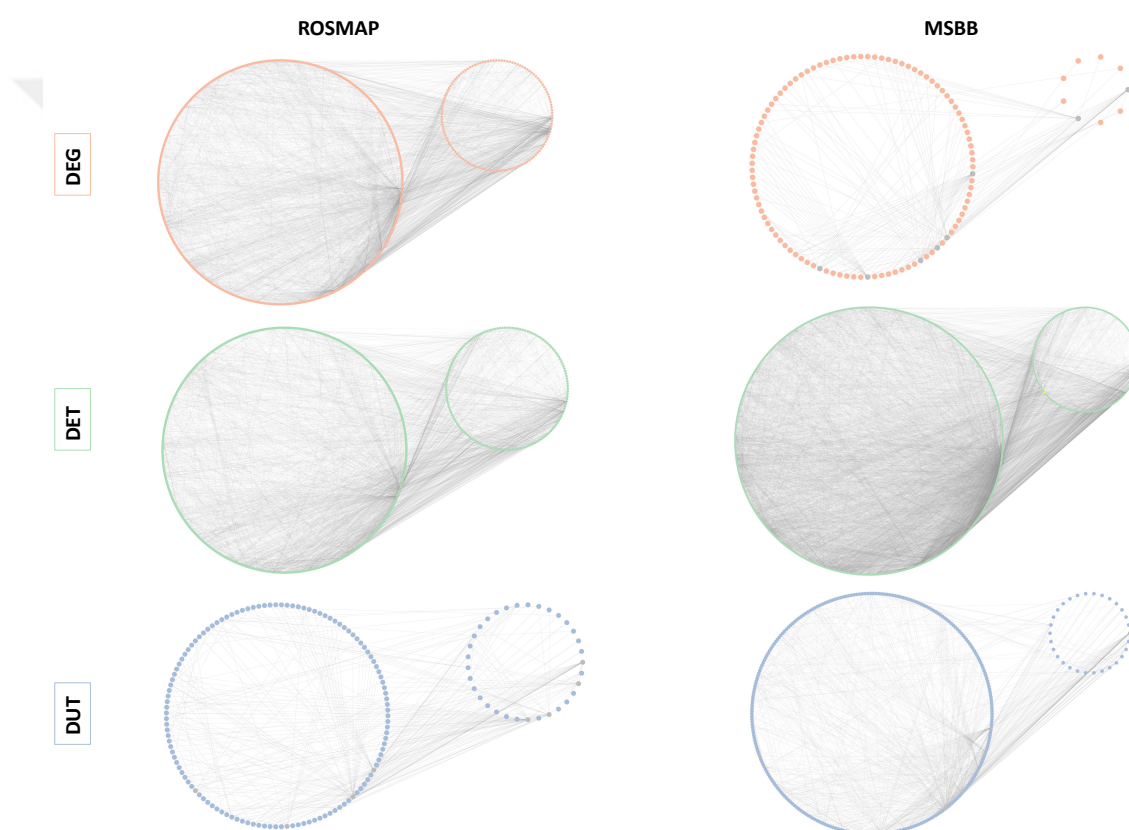


Figure 3.13: The six subnetworks constructed based on the results of the DEG, DET, and DUT analyses from the ROSMAP and MSBB datasets. The small circles on the right of the six subnetworks represent AD-related or LMC-related genes.

The shared genes between the two datasets were discussed in detail in Section 3.2 and analysed through the subnetwork in Figure 3.13. However, capturing the differences between the datasets through the subnetworks is also noteworthy. Certain genes were not common in the two subnetworks but were unique to only one of the datasets, and they have also been mentioned in the previous AD studies. In some cases, AD-related or LMC-related genes from the same protein family were identified across different datasets or analyses such that some appeared in one dataset/analysis while others appeared in the other dataset/analysis. Examples of such genes are the DLGAP1 (DET) gene from the ROSMAP subnetwork and the DLGAP2 (DET) and DLGAP3 (DUT) genes from the MSBB subnetwork. In the article of Seto et al. (2021) on resilient genes and pathways against AD, it was mentioned that inhibiting the phosphorylation of DLGAP1 had a protective effect on synaptic function in the presence of AD pathology. Similarly, high levels of DLGAP2 and other DLGAP family genes were shown to be protective genes for synaptic function and protective genes against AD. Variants of the PICALM (DET) gene, another gene mentioned in that article (Seto et al. 2021), were also stated as involved in pathways including A β clearance, APP processing, synaptic function and neurotransmission. This gene has only been identified in the MSBB subnetwork. Moreover, a genetic variant (rs1990621-G) of the TMEM106B gene has been reported as a neuroprotective factor in individuals with neurodegenerative disorders, including AD, in that study. Notably, this gene is also highlighted in the MSBB subnetwork (DET). Another important gene mentioned in this article is the SORL1 gene, which appeared in both the ROSMAP and MSBB subnetworks as DEG and DET, respectively. Loss of function of the SORL1 gene, which plays a role in lipid homeostasis, was previously associated with AD, and it was thought that it may have been protective against AD, with the gain of function with this variant (Seto et al. 2021).

In the study of Johnson et al. (2018), BIN1 (DET) from the ROSMAP subnetwork and PICALM proteins were reported as AD risk factors and protein isoforms of these genes were associated with alternative exon-exon splicing events (Johnson et al. 2018). The BIN1 gene appeared only in the ROSMAP network, while PICALM was exclusive to the MSBB subnetwork, as mentioned earlier. Again, one of the important genes only appeared in the ROSMAP subnetwork is the CAMKK2 (DET) gene. CAMKK2 is a calcium/calmodulin-dependent protein kinase that plays a role in critical processes

such as synaptic function and neurotransmitter release (Park et al. 2017). This gene is specifically related to CAMK4 and other CAMK genes, which appeared both in the ROSMAP subnetwork as DEG and DET and in the MSBB subnetwork as DET. In the case of the ROSMAP subnetwork, there was an interaction between CAMKK2 and CAMK4. This relationship is based on the ability of CAMKK2 to activate downstream signalling pathways, including CAMK4, which is directly related to memory formation (O'Byrne et al. 2020). In particular, the accumulation of soluble forms of A β was shown to activate the CaMKK-CaMK4 signalling pathway and lead to the phosphorylation of the synapsin, resulting in the disintegration of the synaptic vesicle-synapsin-actin complex (Park et al. 2017). It was predicted that this may adversely affect the formation of new synapses and synaptic plasticity. Therefore, observing the CAMKK2-CAMK4 relationship in the ROSMAP subnetwork may be important to elucidate the potential mechanisms associated with the synaptic dysfunctions in AD.

Similarly, while RAPGEF3 appeared in the ROSMAP subnetwork, RAPGEF4 was in the MSBB subnetwork (Figure 3.14, Figure 3.15). RAPGEF proteins, also known as EPAC (exchange protein directly activated by cAMP) proteins, are GEFs (guanine nucleotide-exchange factors) that activate Rap GTPases when they bind to cAMP (Mcphee et al., n). They were reported to play a role in maintaining LTP processes by converting early LTP into long-lasting LTP. RAPGEF3 (EPAC1) and RAPGEF4 (EPAC2) are genes that play important roles in cAMP-mediated signal transduction and have different effects on AD. EPAC1 regulates neuronal survival and energy metabolism by promoting synaptic plasticity, while EPAC2 plays a critical role in terms of memory regulation, specifically. Studies aiming to silence the EPAC2 gene through gene knock-out or using siRNAs have shown that EPAC2 is more important in memory regulation compared to EPAC1 (Ostroveanu et al. 2010). Therefore, it is suggested that specific modulation of EPAC2 may be a potential therapeutic strategy that may ameliorate impaired memory retrieval in the early stages of AD. EPAC2 activation has been shown to strengthen synaptic plasticity and memory by increasing presynaptic neurotransmitter release and facilitating postsynaptic GluA3 activation. However, its aberrant activation can cause synaptic loss and neuroinflammation. Activation of EPAC1, on the other hand, can improve mitochondrial functions, reducing oxidative stress and slowing down neurodegenerative processes. Balanced

regulation of EPAC1 and EPAC2 by increasing LTP can be considered an important strategy in the treatment of AD (S. Suzuki et al. 2010; T. Zhang and Schmidt 2025).

NDUFS1, NDUFA7, and NDUFA13, which are particularly prominent in the MSBB subnetwork, are subunits of the mitochondrial respiratory Complex I. In AD, mitochondrial dysfunction plays an important role, and changes in the subunits of this complex may contribute to the pathogenesis of the disease. In particular, the levels of NDUFS1 and NDUFA7 proteins were found to be prominently low in LOAD patients (Adav, Park, and Sze 2019). This decrease can lead to mitochondrial dysfunction and consequent disruptions in neuronal energy metabolism. Regarding NDUFA13, biallelic mutations in this gene have been shown to cause mitochondrial Complex I deficiency and neurodevelopmental disorders (Kaiyrzhanov et al. 2025). Although these mutations have not been directly associated with AD, it is thought that disruptions in the function of mitochondrial Complex I may contribute to neurodegenerative processes. As a result, alterations in mitochondrial Complex I subunits such as NDUFS1, NDUFA7, and NDUFA13 have attracted attention as genes that may play a role in the development and progression of AD through mitochondrial dysfunction.

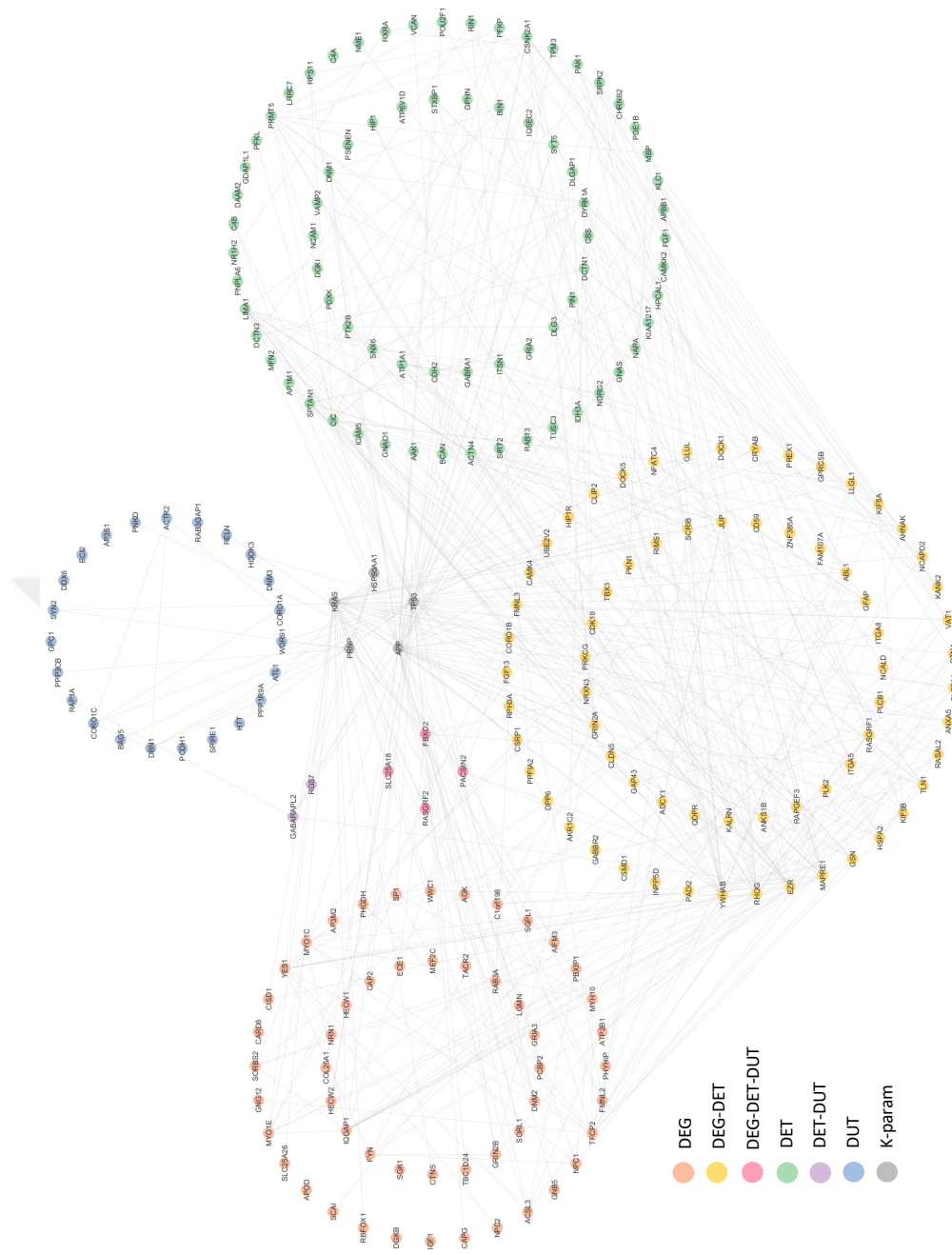


Figure 3.14: Subnetwork results from DEG, DET, and DUT analyses from the ROSMAP dataset. The nodes here represent only AD-related or LMC-related genes.

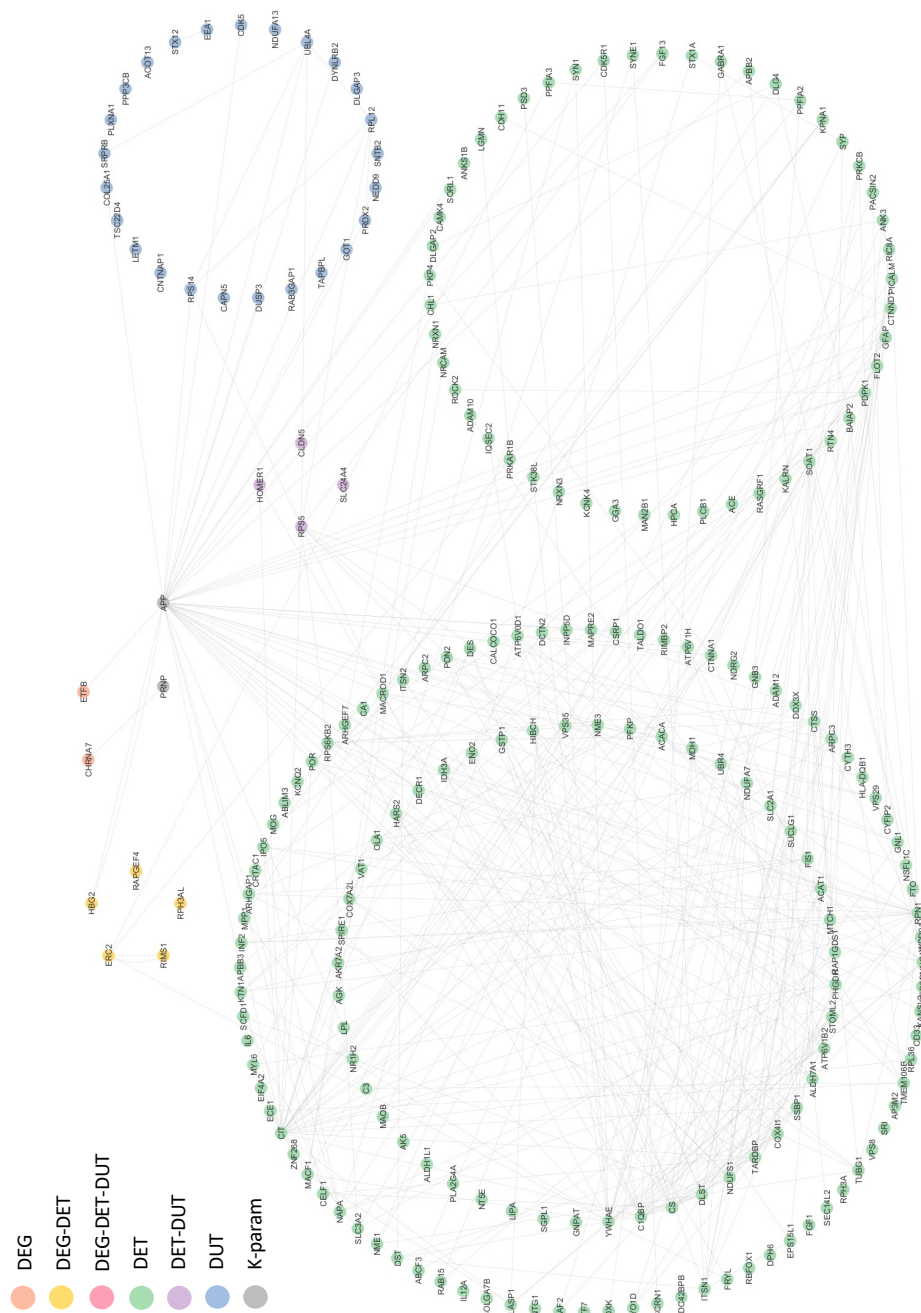


Figure 3.15: Subnetwork results from DEG, DET, and DUT analyses from the MSBB dataset. The nodes here represent only AD-related or LMC-related genes.

3.4. Integration of Two Different Mouse Datasets to Identify and Verify Candidate Genes

In the list of 230 candidate genes, 45 of them showed an intersection with AD-related or LMC-related genes. However, 185 of them may also be important in the progression from AsymAD to AD (Figure 3.4b). Therefore, transcriptome datasets obtained from AD and AsymAD mouse models were analysed in this section. In the first dataset (GSE147434), 876 DEGs were obtained (p-value <0.05) by comparing gene expressions of memory-impaired and memory-intact mice with AD pathology. In the second dataset (GSE111940), 2432 DEGs were found (p-value <0.05) by comparing untreated AD mice with hydroxyurea-treated AD mice, which showed neuroprotection in the AD mice model. The first dataset shared 65 genes from ROSMAP DEGs, 68 genes from ROSMAP DETs, and 16 genes from ROSMAP DUTs. The same dataset shared 10 genes from MSBB DEGs, 129 genes from MSBB DETs, and 37 genes from MSBB DUTs. The second dataset shared 27 genes from ROSMAP DEGs, 30 genes from ROSMAP DETs, 3 genes from ROSMAP DUTs, and 6 genes from MSBB DEGs, 42 genes from MSBB DETs, 6 genes from DUTs. The comparison of candidate genes with the mouse datasets also enabled the discovery of genes that were not available in the AD-related or LMC-related gene lists. These genes can be novel candidate genes for AsymAD. Three of the DEGs in the first mouse dataset matched with 185 genes that are neither AD-related nor LMC-related, as shown in the Venn diagram in Figure 3.4b. These genes are SLC6A12, UXS1, and SPATA13. Nine of the DEGs in the second mouse dataset matched with the 185 genes. These genes are MRPL30, EPHB4, PDK2, OSBPL6, PSME1, WDR19, DCBLD1, FXYP7, and RPL41. From these genes, SLC6A12, OSBPL6, MRPL30, PDK2, EPHB4, and RPL41 genes have been previously mentioned in the context of AD-related studies.

Results from the integration of mouse datasets indicated that SLC6A12, OSBPL6, MRPL30, PDK2, EPHB4, PSME1, and RPL41 genes were previously mentioned in the literature in relation to AD. Among those genes, MRPL30 and RPL41 were already discussed above. SLC6A12 is a subfamily of gamma-aminobutyric acid (GABA), and the upregulation of this gene accompanying the downregulation of GABA signalling has been associated with cognitive impairment in AD (Zhao, Yao, and Li 2021). OSBPL6 was reported to carry variants that contributed to a genetic predisposition to AD and was identified as a gene linked with AD risk (Herold et al. 2016). PDK2 has

attracted attention as an enzyme involved in astrocyte-mediated neuroinflammation and metabolic disorders related to AD, but detailed studies directly linking it to AD are limited (Rahman and Suk 2020). EPHB4 is a member of the EPH family. One of its family genes, EPHA1, was associated with neurodegenerative diseases, and it was shown that its genetic variations may affect AD susceptibility (Hollingworth et al. 2011). In another study, overexpression of EphB2 was observed to protect neurons from A β induced toxicity by regulating NMDA receptor signalling (Geng et al. 2013).



4. CONCLUSIONS

The findings of this thesis study provided new insights into the molecular mechanisms underlying AsymAD vs AD by integrating DEG, DET, and DUT analyses in two well-established AD datasets. Several candidate genes that may play critical roles in synaptic function, ECM integrity, and mitochondrial energy metabolism were identified. In particular, NRXN3 and DGKB were highlighted as potential biomarker genes due to their involvement in synaptic plasticity and neurotransmitter release and their altered expression patterns in AD and AsymAD. Additionally, the role of ADAMTS2 in ECM dysregulation and its association with cognitive decline were highlighted, along with the importance of certain collagen-related genes (PCOLCE, COL25A1, COL26A1, ADAMTS2) in maintaining ECM integrity. New findings on GNG4 and its regulation at the transcript level suggest its role in synapse-related molecular mechanisms in AD progression. Furthermore, the increased ratio of ENPP5 isoform in AsymAD may point to early neuroprotective or compensatory mechanisms associated with NAD-based neurotransmission. Finally, dysregulation of mitochondrial ribosomal genes such as MRPL1 and MRPL30, critical for energy metabolism, further highlights the interplay of mitochondrial dysfunction in AD pathology. In addition, a striking mechanism in the comparison of AD vs AsymAD is synaptic plasticity, which plays a role in memory formation and the LTP mechanism, where synapses are strengthened due to active use. These may be the most extensive mechanisms involved in the protection of AsymAD individuals from cognitive decline. Thus, despite the presence of pathology in AsymAD individuals, the main reason for the absence of cognitive impairment could be attributed to the strong synaptic plasticity and signal transduction between synapses.

A limitation of this study is the inconsistencies observed across the two studied datasets in both the number of transcripts and their regulation levels due to differences between the datasets and the analytical methods used. As a suggestion for future directions, it is envisaged that cellular heterogeneity can be examined in detail by using scRNA-seq AD datasets after identifying individuals with AsymAD using neuropathological and clinical scores, as presented in this thesis study. Single-cell-level analyses can help us understand which cell groups are more affected in AD and AsymAD states or which pathways are activated specifically for cell types by revealing

transcriptomic differences between the cell types. Such studies may offer the possibility to more precisely identify the molecular changes that occur in the early stages of the disease. In addition, AsymAD mouse models can be used for the validation of the candidate genes identified in this study.

These results collectively highlight the importance of integrating multi-layered transcriptomic analyses to uncover key molecular pathways and potential therapeutic targets in AD and AsymAD.



REFERENCES

- Adav, Sunil S., Jung Eun Park, and Siu Kwan Sze. 2019. "Quantitative Profiling Brain Proteomes Revealed Mitochondrial Dysfunction in Alzheimer's Disease." *Molecular Brain* 12(1). doi:10.1186/s13041-019-0430-y.
- Anderson, Nicole D. 2019. "State of the Science on Mild Cognitive Impairment (MCI)." *CNS Spectrums* 24(1): 78–87. doi:10.1017/S1092852918001347.
- Atri, Alireza. 2019. "The Alzheimer's Disease Clinical Spectrum: Diagnosis and Management." *Medical Clinics of North America* 103(2): 263–93. doi:10.1016/j.mcna.2018.10.009.
- Bader, Hannah L., Douglas R. Keene, Benjamin Charvet, Guido Veit, Wolfgang Driever, Manuel Koch, and Florence Ruggiero. 2009. "Zebrafish Collagen XII Is Present in Embryonic Connective Tissue Sheaths (Fascia) and Basement Membranes." *Matrix Biology* 28(1): 32–43. doi:10.1016/j.matbio.2008.09.580.
- Bayraktar, Abdulahad, Simon Lam, Ozlem Altay, Xiangyu Li, Meng Yuan, Cheng Zhang, Muhammad Arif, et al. 2021. "Revealing the Molecular Mechanisms of Alzheimer's Disease Based on Network Analysis." *International Journal of Molecular Sciences* 22(21). doi:10.3390/ijms222111556.
- Van Den Berge, Koen, Katharina M Hembach, Charlotte Soneson, Simone Tiberi, Lieven Clement, Michael I Love, Rob Patro, and Mark D Robinson. 2025. "Annual Review of Biomedical Data Science RNA Sequencing Data: Hitchhiker's Guide to Expression Analysis INTRODUCTION: OVERVIEW OF THE RNA SEQUENCING ASSAY." doi:10.1146/annurev-biodatasci-072018.
- Blacher, Eran, Tulin Dadali, Alina Bespalko, Viola J. Haupenthal, Marcus O.W. Grimm, Tobias Hartmann, Frances E. Lund, Reuven Stein, and Ayelet Levy. 2015. "Alzheimer's Disease Pathology Is Attenuated in a CD38-Deficient Mouse Model." *Annals of Neurology* 78(1): 88–103. doi:10.1002/ana.24425.
- Bonham, Luke W., Daniel S. Evans, Yongmei Liu, Steven R. Cummings, Kristine Yaffe, and Jennifer S. Yokoyama. 2018. "Neurotransmitter Pathway Genes in Cognitive Decline During Aging: Evidence for GNG4 and KCNQ2 Genes." *American Journal of Alzheimer's Disease and other Dementias* 33(3): 153–65. doi:10.1177/1533317517739384.
- Bot, Nathalie, Claude Schweizer, Saoussen Ben Halima, and Patrick C. Fraering. 2011. "Processing of the Synaptic Cell Adhesion Molecule Neurexin-3 β by Alzheimer Disease α - and γ -Secretases." *Journal of Biological Chemistry* 286(4): 2762–73. doi:10.1074/jbc.M110.142521.
- Bouter, Yvonne, Tim Kacprowski, Robert Weissmann, Katharina Dietrich, Henning Borgers, Andreas Brauß, Christian Sperling, et al. 2014. "Deciphering the Molecular Profile of Plaques, Memory Decline and Neuron Loss in Two Mouse Models for Alzheimer's Disease by Deep Sequencing." *Frontiers in Aging Neuroscience* 6(APR). doi:10.3389/fnagi.2014.00075.

Brose, Rebecca Deering, Elin Lehrmann, Yongqing Zhang, Roger H. Reeves, Kirby D. Smith, and Mark P. Mattson. 2018. "Hydroxyurea Attenuates Oxidative, Metabolic, and Excitotoxic Stress in Rat Hippocampal Neurons and Improves Spatial Memory in a Mouse Model of Alzheimer's Disease." *Neurobiology of Aging* 72: 121–33. doi:10.1016/j.neurobiolaging.2018.08.021.

Busche, Marc Aurel, and Bradley T. Hyman. 2020. "Synergy between Amyloid- β and Tau in Alzheimer's Disease." *Nature Neuroscience* 23(10): 1183–93. doi:10.1038/s41593-020-0687-6.

Chaubey, Kalyani, Edwin Vázquez-Rosa, Min-Kyoo Shin, Matasha Dhar, Kathryn Franke, Coral Cintrón-Pérez, Yeojung Koh, et al. 2022. "Mechanisms of Protection in Early and Mid-Stage Alzheimer's Disease by Treatment with NAD⁺ + Preserving Compound." *Alzheimer's & Dementia* 18(S3). doi:10.1002/alz.067922.

Ciurea, Alexandru Vlad, Aurel George Mohan, Razvan Adrian Covache-Busioc, Horia Petre Costin, Luca Andrei Glavan, Antonio Daniel Corlatescu, and Vicentiu Mircea Saceleanu. 2023. "Unraveling Molecular and Genetic Insights into Neurodegenerative Diseases: Advances in Understanding Alzheimer's, Parkinson's, and Huntington's Diseases and Amyotrophic Lateral Sclerosis." *International Journal of Molecular Sciences* 24(13). doi:10.3390/ijms241310809.

Deture, Michael A., and Dennis W. Dickson. 2019. "The Neuropathological Diagnosis of Alzheimer's Disease." *Molecular Neurodegeneration* 14(1). doi:10.1186/s13024-019-0333-5.

Dick, Fiona, Gonzalo S. Nido, Guido Werner Alves, Ole Bjørn Tysnes, Gry Hilde Nilsen, Christian Dölle, and Charalampos Tzoulis. 2020. "Differential Transcript Usage in the Parkinson's Disease Brain." *PLoS Genetics* 16(11). doi:10.1371/journal.pgen.1009182.

Driscoll, I, and J Troncoso. 2011. *Asymptomatic Alzheimer's Disease: A Prodrome or a State of Resilience?*

Düz, Elif, and Tunahan Çakır. 2024. "Effect of RNA-Seq Data Normalization on Protein Interactome Mapping for Alzheimer's Disease." *Computational Biology and Chemistry* 109. doi:10.1016/j.compbiolchem.2024.108028.

Fan, Liyuan, Chengyuan Mao, Xinchao Hu, Shuo Zhang, Zhihua Yang, Zhengwei Hu, Huifang Sun, et al. 2020. "New Insights Into the Pathogenesis of Alzheimer's Disease." *Frontiers in Neurology* 10. doi:10.3389/fneur.2019.01312.

Farhadieh, Mohammad Erfan, and Kamran Ghaedi. 2023. "Analyzing Alternative Splicing in Alzheimer's Disease Postmortem Brain: A Cell-Level Perspective." *Frontiers in Molecular Neuroscience* 16. doi:10.3389/fnmol.2023.1237874.

Geng, Dandan, Lin Kang, Yuhong Su, Jianxin Jia, Jun Ma, Sha Li, Juan Du, and Huixian Cui. 2013. "Protective Effects of EphB2 on A β 1-42 Oligomer-Induced Neurotoxicity and Synaptic NMDA Receptor Signaling in Hippocampal Neurons." *Neurochemistry International* 63(4): 283–90. doi:10.1016/j.neuint.2013.06.016.

Gentry, Charisse, Michael Malek-Ahmadi, Susan Bolas, and Jose Pena. 2024. "PET Imaging in Alzheimer Disease: Pathology and Research Insights for Technologists." *Journal of nuclear medicine technology* 52(4): 306–11. doi:10.2967/jnmt.124.268916.

Gong, Bing, Yong Pan, Prashant Vempati, Wei Zhao, Lindsay Knable, Lap Ho, Jun Wang, et al. 2013. "Nicotinamide Riboside Restores Cognition through an Upregulation of Proliferator-Activated Receptor- γ Coactivator 1 α Regulated β -Secretase 1 Degradation and Mitochondrial Gene Expression in Alzheimer's Mouse Models." *Neurobiology of Aging* 34(6): 1581–88. doi:10.1016/j.neurobiolaging.2012.12.005.

Gurses, Murat Serdar, Mustafa Numan Ural, Mehmet Akif Gulec, Omer Akyol, and Sumeyya Akyol. 2016. "Pathophysiological Function of ADAMTS Enzymes on Molecular Mechanism of Alzheimer's Disease." *Aging and Disease* 7(4): 479–90. doi:10.14336/AD.2016.0111.

Hampel, Harald, Robert Nisticò, Nicholas T. Seyfried, Allan I. Levey, Erica Modeste, Pablo Lemercier, Filippo Baldacci, et al. 2021. "Omics Sciences for Systems Biology in Alzheimer's Disease: State-of-the-Art of the Evidence." *Ageing Research Reviews* 69. doi:10.1016/j.arr.2021.101346.

Herold, C., B. V. Hooli, K. Mullin, T. Liu, J. T. Roehr, M. Mattheisen, A. R. Parrado, et al. 2016. "Family-Based Association Analyses of Imputed Genotypes Reveal Genome-Wide Significant Association of Alzheimer's Disease with OSBPL6, PTPRG, and PDCL3." *Molecular Psychiatry* 21(11): 1608–12. doi:10.1038/mp.2015.218.

Hollingworth, Paul, Denise Harold, Rebecca Sims, Amy Gerrish, Jean Charles Lambert, Minerva M. Carrasquillo, Richard Abraham, et al. 2011. "Common Variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP Are Associated with Alzheimer's Disease." *Nature Genetics* 43(5): 429–36. doi:10.1038/ng.803.

Hozumi, Yasukazu, and Kaoru Goto. 2012. "Diacylglycerol Kinase β in Neurons: Functional Implications at the Synapse and in Disease." *Advances in Biological Regulation* 52(2): 315–25. doi:10.1016/j.jbior.2012.03.003.

Jellinger, Kurt A, and K A Jellinger. *Clinical Validity of Braak Staging in the Oldest-Old EDITORIAL NOTE*.

Johnson, Erik C.B., Eric B. Dammer, Duc M. Duong, Luming Yin, Madhav Thambisetty, Juan C. Troncoso, James J. Lah, Allan I. Levey, and Nicholas T. Seyfried. 2018. "Deep Proteomic Network Analysis of Alzheimer's Disease Brain Reveals Alterations in RNA Binding Proteins and RNA Splicing Associated with Disease." *Molecular Neurodegeneration* 13(1). doi:10.1186/s13024-018-0282-4.

Juganavar, Anup, Abhishek Joshi, and Tejas Shegekar. 2023. "Navigating Early Alzheimer's Diagnosis: A Comprehensive Review of Diagnostic Innovations." *Cureus*. doi:10.7759/cureus.44937.

Kaiyrzhanov, Rauan, Kyle Thompson, Stephanie Efthymiou, Askhat Mukushev, Akbota Zharylkassyn, Chitra Prasad, Ehsan Ghayoor Karimiani, et al. 2025. "Biallelic

NDUFA13 Variants Lead to a Neurodevelopmental Phenotype with Gradual Neurological Impairment.” *Brain Communications* 7(1). doi:10.1093/braincomms/fcae453.

Knopman, David S., Helene Amieva, Ronald C. Petersen, Gäel Chételat, David M. Holtzman, Bradley T. Hyman, Ralph A. Nixon, and David T. Jones. 2021. “Alzheimer Disease.” *Nature Reviews Disease Primers* 7(1). doi:10.1038/s41572-021-00269-y.

Kong, Wei, Xiaoyang Mou, Na Zhang, Weiming Zeng, Shasha Li, and Yang Yang. 2015. “The Construction of Common and Specific Significance Subnetworks of Alzheimer’s Disease from Multiple Brain Regions.” *BioMed Research International* 2015. doi:10.1155/2015/394260.

Lee, Dongwon, Eunjoon Kim, and Keiko Tanaka-Yamamoto. 2016. “Diacylglycerol Kinases in the Coordination of Synaptic Plasticity.” *Frontiers in Cell and Developmental Biology* 4(AUG). doi:10.3389/fcell.2016.00092.

Lepelletier, F. X., D. M.A. Mann, A. C. Robinson, E. Pinteaux, and H. Boutin. 2017. “Early Changes in Extracellular Matrix in Alzheimer’s Disease.” *Neuropathology and Applied Neurobiology* 43(2): 167–82. doi:10.1111/nan.12295.

Li, Yijie (Jamie), Elizabeth Kresock, Rayus Kuplicki, Jonathan Savitz, and Brett A. McKinney. 2022. “Differential Expression of MDGA1 in Major Depressive Disorder.” *Brain, Behavior, and Immunity - Health* 26. doi:10.1016/j.bbih.2022.100534.

Lunnon, Katie, Aoife Keohane, Ruth Pidsley, Stephen Newhouse, Joanna Riddoch-Contreras, Elisabeth B. Thubron, Matthew Devall, et al. 2017. “Mitochondrial Genes Are Altered in Blood Early in Alzheimer’s Disease.” *Neurobiology of Aging* 53: 36–47. doi:10.1016/j.neurobiolaging.2016.12.029.

Marques-Coelho, Diego, Lukas da Cruz Carvalho Iohan, Ana Raquel Melo de Farias, Amandine Flaig, Franck Letournel, Marie Laure Martin-Négrier, Françoise Chapon, et al. 2021. “Differential Transcript Usage Unravels Gene Expression Alterations in Alzheimer’s Disease Human Brains.” *npj Aging and Mechanisms of Disease* 7(1). doi:10.1038/s41514-020-00052-5.

McCorkindale, Andrew N., Ellis Patrick, James A. Duce, Boris Guennewig, and Greg T. Sutherland. 2022. “The Key Factors Predicting Dementia in Individuals With Alzheimer’s Disease-Type Pathology.” *Frontiers in Aging Neuroscience* 14. doi:10.3389/fnagi.2022.831967.

Mcphee, I, L C D Gibson, J Kewney, C Darroch, P A Stevens, D Spinks, A Cooreman, and S J Mackenzie. *Cyclic Nucleotide Signalling: A Molecular Approach to Drug Discovery for Alzheimer’s Disease*.

Mei, Taoyu, Yuan Li, Anna Orduña Dolado, Zhiquan Li, Robin Andersson, Laura Berliocchi, and Lene Juel Rasmussen. 2023. “Pooled Analysis of Frontal Lobe Transcriptomic Data Identifies Key Mitophagy Gene Changes in Alzheimer’s Disease Brain.” *Frontiers in Aging Neuroscience* 15. doi:10.3389/fnagi.2023.1101216.

Monsell, Sarah E, Charles Mock, Jason Hassenstab, Catherine M Roe, Nigel J Cairns, John C Morris, and Walter Kukull. 2014. *Neuropsychological Changes in Asymptomatic Persons with Alzheimer Disease Neuropathology*.

Morris, John C. 1997. "Clinical Dementia Rating: A Reliable and Valid Diagnostic and Staging Measure for Dementia of the Alzheimer Type." In *International Psychogeriatrics*, , 173–76. doi:10.1017/S1041610297004870.

Nakamura, Akinori, Pablo Cuesta, Takashi Kato, Yutaka Arahata, Kaori Iwata, Misako Yamagishi, Izumi Kuratsubo, et al. 2017. "Early Functional Network Alterations in Asymptomatic Elders at Risk for Alzheimer's Disease." *Scientific Reports* 7(1). doi:10.1038/s41598-017-06876-8.

Nguyen, Hai Duc, Woong Ki Kim, and Huong Huong Vu. 2024. "Molecular Mechanisms Implicated in Protein Changes in the Alzheimer's Disease Human Hippocampus." *Mechanisms of Ageing and Development* 219. doi:10.1016/j.mad.2024.111930.

O'Byrne, Sean N., John W. Scott, Joseph R. Pilotte, André S. da Santiago, Christopher G. Langendorf, Jonathan S. Oakhill, Benjamin J. Eduful, et al. 2020. "In Depth Analysis of Kinase Cross Screening Data to Identify CaMKK2 Inhibitory Scaffolds." *Molecules* 25(2). doi:10.3390/molecules25020325.

Ostroveanu, Anghelus, Eddy A. Van Der Zee, Ulrich L.M. Eisel, Martina Schmidt, and Ingrid M. Nijholt. 2010. "Exchange Protein Activated by Cyclic AMP 2 (Epac2) Plays a Specific and Time-Limited Role in Memory Retrieval." *Hippocampus* 20(9): 1018–26. doi:10.1002/hipo.20700.

Park, Daehun, Myeongsu Na, Jung Ah Kim, Unghwi Lee, Eunji Cho, Mirye Jang, and Sunghoe Chang. 2017. *Activation of CaMKIV by Soluble Amyloid-b 1-42 Impedes Trafficking of Axonal Vesicles and Impairs Activity-Dependent Synaptogenesis*. <http://stke.sciencemag.org/>.

Patel, Hamel, Angela K. Hodges, Charles Curtis, Sang Hyuck Lee, Claire Troakes, Richard J.B. Dobson, and Stephen J. Newhouse. 2019. "Transcriptomic Analysis of Probable Asymptomatic and Symptomatic Alzheimer Brains." *Brain, Behavior, and Immunity* 80: 644–56. doi:10.1016/j.bbi.2019.05.009.

Pérez-González, Marta, Maite Mendioroz, Sara Badesso, Diego Sucunza, Miren Roldan, Maria Espelosín, Susana Ursua, et al. 2020. "PLA2G4E, a Candidate Gene for Resilience in Alzheimer's Disease and a New Target for Dementia Treatment." *Progress in Neurobiology* 191. doi:10.1016/j.pneurobio.2020.101818.

Prokopec, Stephenie D., John D. Watson, Daryl M. Waggott, Ashley B. Smith, Alexander H. Wu, Allan B. Okey, Raimo Pohjanvirta, and Paul C. Boutros. 2013. "Systematic Evaluation of Medium-Throughput mRNA Abundance Platforms." *RNA* 19(1): 51–62. doi:10.1261/rna.034710.112.

Puthiyedth, Nisha, Carlos Riveros, Regina Berretta, and Pablo Moscato. 2016. "Identification of Differentially Expressed Genes through Integrated Study of

Alzheimer's Disease Affected Brain Regions." *PLoS ONE* 11(4). doi:10.1371/journal.pone.0152342.

Rahman, Md Habibur, and Kyoungso Suk. 2020. "Mitochondrial Dynamics and Bioenergetic Alteration During Inflammatory Activation of Astrocytes." *Frontiers in Aging Neuroscience* 12. doi:10.3389/fnagi.2020.614410.

Robert, M. 2013. "Epidemiology and Clinical Diagnosis: Alzheimer Disease." *PET Clinics* 8(4): 391–405. doi:10.1016/j.cpet.2013.08.001.

Rose, Keron W.J., Nandaraj Taye, Stylianos Z. Karoulias, and Dirk Hubmacher. 2021. "Regulation of ADAMTS Proteases." *Frontiers in Molecular Biosciences* 8. doi:10.3389/fmolb.2021.701959.

Rossetti, Heidi C, C Munro Cullum, Linda S Hynan, and Laura H Lacritz. 24 Alzheimer Dis Assoc Disord *The CERAD Neuropsychologic Battery Total Score and the Progression of Alzheimer Disease*. www.alzheimerjournal.com.

Seto, Mabel, Rebecca L. Weiner, Logan Dumitrescu, and Timothy J. Hohman. 2021. "Protective Genes and Pathways in Alzheimer's Disease: Moving towards Precision Interventions." *Molecular Neurodegeneration* 16(1). doi:10.1186/s13024-021-00452-5.

Seyfried, Nicholas T., Eric B. Dammer, Vivek Swarup, Divya Nandakumar, Duc M. Duong, Luming Yin, Qidong Deng, et al. 2017. "A Multi-Network Approach Identifies Protein-Specific Co-Expression in Asymptomatic and Symptomatic Alzheimer's Disease." *Cell Systems* 4(1): 60-72.e4. doi:10.1016/j.cels.2016.11.006.

Smagin, Dmitry A., Anna G. Galyamina, Irina L. Kovalenko, Vladimir N. Babenko, and Natalia N. Kudryavtseva. 2019. "Aberrant Expression of Collagen Gene Family in the Brain Regions of Male Mice with Behavioral Psychopathologies Induced by Chronic Agonistic Interactions." *BioMed Research International* 2019. doi:10.1155/2019/7276389.

Sommer, Markus J., Sooyoung Cha, Ales Varabyou, Natalia Rincon, Sukhwan Park, Ilia Minkin, Mihaela Perte, Martin Steinegger, and Steven L. Salzberg. 2022. "Structure-Guided Isoform Identification for the Human Transcriptome." *eLife* 11. doi:10.7554/ELIFE.82556.

Steiglitz, Barry M., Jaclynn M. Kreider, Elizabeth P. Frankenburg, William N. Pappano, Guy G. Hoffman, Jeffrey A. Meganck, Xiaowen Liang, et al. 2006. "Procollagen C Proteinase Enhancer 1 Genes Are Important Determinants of the Mechanical Properties and Geometry of Bone and the Ultrastructure of Connective Tissues." *Molecular and Cellular Biology* 26(1): 238–49. doi:10.1128/mcb.26.1.238-249.2006.

Sun, Yahan, Sen Xu, Ming Jiang, Xia Liu, Liang Yang, Zhantao Bai, and Qinghu Yang. 2021. "Role of the Extracellular Matrix in Alzheimer's Disease." *Frontiers in Aging Neuroscience* 13. doi:10.3389/fnagi.2021.707466.

Suzuki, Masayoshi, Kenta Tezuka, Takumi Handa, Risa Sato, Hina Takeuchi, Masaki Takao, Mitsutoshi Tano, and Yasuo Uchida. 2022. "Upregulation of Ribosome

Complexes at the Blood-Brain Barrier in Alzheimer's Disease Patients." *Journal of Cerebral Blood Flow and Metabolism* 42(11): 2134–50. doi:10.1177/0271678X221111602.

Suzuki, Sayaka, Utako Yokoyama, Takaya Abe, Hiroshi Kiyonari, Naoya Yamashita, Yuko Kato, Reiko Kurotani, et al. 2010. "Differential Roles of Epac in Regulating Cell Death in Neuronal and Myocardial Cells." *Journal of Biological Chemistry* 285(31): 24248–59. doi:10.1074/jbc.M109.094581.

Swarup, Vivek, Timothy S. Chang, Duc M. Duong, Eric B. Dammer, Jingting Dai, James J. Lah, Erik C.B. Johnson, et al. 2020. "Identification of Conserved Proteomic Networks in Neurodegenerative Dementia." *Cell Reports* 31(12). doi:10.1016/j.celrep.2020.107807.

Thal, Dietmar R, Udo Rüb, Mario Orantes, and Heiko Braak. 2002. *Phases of A-Deposition in the Human Brain and Its Relevance for the Development of AD*. www.neurology.org.

Tong, Ying, Ying Xu, Kimberly Searce-Levie, Louis J. Ptáček, and Ying Hui Fu. 2010. "COL25A1 Triggers and Promotes Alzheimer's Disease-like Pathology in Vivo." *Neurogenetics* 11(1): 41–52. doi:10.1007/s10048-009-0201-5.

Vu, Megan, Rohan Mangal, Thor Stead, Carlos Lopez-Ortiz, and Latha Ganti. 2022. "Impact of Alzheimer's Disease on Caregivers in the United States." *Health Psychology Research* 10(3). doi:10.52965/001c.37454.

Wang, Qi, Kewei Chen, Yi Su, Eric M. Reiman, Joel T. Dudley, and Benjamin Readhead. 2022. "Deep Learning-Based Brain Transcriptomic Signatures Associated with the Neuropathological and Clinical Severity of Alzheimer's Disease." *Brain Communications* 4(1). doi:10.1093/braincomms/fcab293.

WANG, YAN, JING LIU, BO HUANG, YAN-MEI XU, JING LI, LIN-FENG HUANG, JIN LIN, et al. 2015. "Mechanism of Alternative Splicing and Its Regulation." *Biomedical Reports* 3(2): 152–58. doi:10.3892/br.2014.407.

Williams, Jamal B., Qing Cao, and Zhen Yan. 2021. "Transcriptomic Analysis of Human Brains with Alzheimer's Disease Reveals the Altered Expression of Synaptic Genes Linked to Cognitive Deficits." *Brain Communications* 3(3). doi:10.1093/braincomms/fcab123.

Wu, Hao, Jiao Wang, Xiaoyuan Hu, Cheng Zhuang, Jianxin Zhou, Peiru Wu, Shengli Li, and Robert Chunhua Zhao. 2023. "Comprehensive Transcript-Level Analysis Reveals Transcriptional Reprogramming during the Progression of Alzheimer's Disease." *Frontiers in Aging Neuroscience* 15. doi:10.3389/fnagi.2023.1191680.

Yamakage, Yuko, Michinao Kato, Aya Hongo, Himari Ogino, Keisuke Ishii, Takumi Ishizuka, Takana Kamei, et al. 2019. "A Disintegrin and Metalloproteinase with Thrombospondin Motifs 2 Cleaves and Inactivates Reelin in the Postnatal Cerebral Cortex and Hippocampus, but Not in the Cerebellum." *Molecular and Cellular Neuroscience* 100. doi:10.1016/j.mcn.2019.103401.

Zhang, Tong, and Martina Schmidt. 2025. "Targeting Epac2 and GluA3-Containing AMPARs: A Novel Therapeutic Strategy for Alzheimer's Disease." *Neural Regeneration Research* 20(11): 3223–24. doi:10.4103/NRR.NRR-D-24-00751.

Zhang, Yida, Harry Taegyun Yang, Kathryn Kadash-Edmondson, Yang Pan, Zhicheng Pan, Beverly L. Davidson, and Yi Xing. 2020. "Regional Variation of Splicing QTLs in Human Brain." *American Journal of Human Genetics* 107(2): 196–210. doi:10.1016/j.ajhg.2020.06.002.

Zhao, Xingxing, Hongmei Yao, and Xinyi Li. 2021. "Unearthing of Key Genes Driving the Pathogenesis of Alzheimer's Disease via Bioinformatics." *Frontiers in Genetics* 12. doi:10.3389/fgene.2021.641100.

Zheng, Jun Juan, Wen Xing Li, Jia Qian Liu, Yi Cheng Guo, Qian Wang, Gong Hua Li, Shao Xing Dai, and Jing Fei Huang. 2018. "Low Expression of Aging-Related NRXN3 Is Associated with Alzheimer Disease: A Systematic Review and Meta-Analysis." *Medicine (United States)* 97(28). doi:10.1097/MD.00000000000011343.

BIOGRAPHY

Rümeysa AKSU got her B.Sc. degree in Molecular Biology and Genetics from İstanbul Arel University in 2021. After a half-year gap, she started her M.Sc. in Bioinformatics and Systems Biology program at the Bioengineering Department of Gebze Technical University in 2022. Additionally, she was a scholarship holder in a TUBITAK-1002-A Project during the dissertation stage of her M.Sc.



PUBLICATIONS AND PRESENTATIONS FROM THE THESIS

Aksu R., Lüleci H. B., Çakır T., (2024), “Geniş Çaplı Transkriptom ve Ağ Analizleriyle Asemptomatik Alzheimer Hastalığına Özgü Moleküler Mekanizmaların Keşfi”, ISCB-SC RSG Türkiye 11th Student Symposium 2024, Pendik, İstanbul, 15-16 November.

Aksu R., Lüleci H. B., Çakır T., (2024), “An RNA-Seq Based Network Approach to Elucidate Molecular Mechanisms of Asymptomatic Alzheimer’s Disease”, The 17th International Symposium on Health Informatics and Bioinformatics (HIBIT) 2024, Fatih, İstanbul, 18-20 December.

