

Transcriptomics in Neurodegenerative Diseases

Kosar Hooshmand



THE UNIVERSITY OF
SYDNEY

A thesis submitted to fulfil requirements for the degree of:
Doctor of Philosophy (PhD)

Supervisor:
Dr. Boris Guennewig

Brain and Mind Centre, Faculty of Medicine, and Health,
School of Medical sciences,
The University of Sydney
December 2022

Statement of originality

I, Kosar Hooshmand¹, declare that this thesis is submitted in fulfilment of the requirements for the conferral of the degree Doctor of Philosophy (PhD), from the University of Sydney. The experiments were carried out by the author, unless stated otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

¹ Kosar Hooshmand, Doctor of Philosophy, December 2022

Publications

Publications pertaining to this thesis

- 1- Hooshmand, K., Halliday, G.M., Pineda, S.S., Sutherland, G.T. and Guennewig, B. Overlap between Central and Peripheral Transcriptomes in Parkinson's Disease but Not Alzheimer's Disease. 2022. International Journal of Molecular Sciences. 23(9), p.5200.
- 2- Hooshmand, K., Halliday, G.M., Pineda, S.S., Sutherland, G.T. and Guennewig, B. RNA-seq based genetic variant discovery from peripheral and central tissues to characterise Alzheimer's and Parkinson's disease (*Submitted or Manuscript in preparation*).
- 3- Hooshmand, K., Halliday, G.M., Pineda, S.S., Sutherland, G.T. and Guennewig, B. Exploring the effect of SNP on gene expression and alternative splicing patterns in blood and brain from Alzheimer's, and Parkinson's Disease (*Submitted or Manuscript in preparation*).

Publications contributed to during candidature

- 1- Theunissen, F., West, P.K., Brennan, S., Petrović, B., Hooshmand, K., Akkari, P.A., Keon, M. and Guennewig, B. New perspectives on cytoskeletal dysregulation and mitochondrial mislocalization in amyotrophic lateral sclerosis. 2021. Translational neurodegeneration. 10(1), pp.1-16.
- 2- Hooshmand, K., Goldstein, D., Timmins, H.C., Li, T., Harrison, M., Friedlander, M.L., Lewis, C.R., Lees, J.G., Moalem-Taylor, G., Guennewig, B. and Park, S.B., 2022. Polygenic risk of paclitaxel-induced peripheral neuropathy: a genome-wide association study. Journal of Translational Medicine, 20(1), pp.1-12.
- 3- Hooshmand, K., Halliday, G.M., Pineda, S.S., Petrovic B., Sutherland, G.T. and Guennewig, B. Nutrition, Epigenetics and Cognition: A balancing act between neuroprotection and neurodegeneration across the lifespan (*Submitted or Manuscript in preparation*).

Conference proceedings

- 1- Sydney Bioinformatics Research Symposium (SBRS). The Sydney Nanoscience Hub, University of Sydney, Australia, June 24, 2022. Overlap between Central and Peripheral Transcriptomes in Parkinson's Disease but Not Alzheimer's Disease.
- 2- AMSI BioInfoSummer 2019. Charles Perkins Centre, The University of Sydney, Sydney, Australia, December 2-6, 2019. Transcriptomics in Neurodegenerative Disease.

Acknowledgements

This thesis could not have been accomplished without the support of many people. First and foremost, I would like to express my sincere gratitude to my primary supervisor *Dr Boris Guennewig*. Your supervision, guidance, encouragement, support, and trust throughout my PhD journey are very much appreciated.

I am also deeply grateful to my co-supervisor *Associated Professor Greg T. Sutherland*, who guided me in the field scientific explanations, improvement of my writing skills, and for the proof-reading.

Thank you also to *Dr Sandy S. Pineda* for your kind and helpful advice with statistical methodologies and guidance with R.

Thank you, *Professor Glenda M. Halliday*, for your support throughout this project, and your life advice.

And a special thanks to *Mr. Matt Keon* for allowing me the opportunity to work at GenieUs Genomics as Bioinformatician & platform consultant, which led to a growth in my bioinformatics skills and capabilities. I really appreciate your understanding and support through my challenging time.

To my BMC friends –*Vivian, Diba, Nazanin, Ali, Susan, and Parisa*, thank you for the continuous support that made this PhD so much more enjoyable. I am glad to have gone through this journey with so many kind and supportive like-minded people.

Finally, a special thank you to those closest to me. To my Parents, for their unconditional love, support, and encouragement through my educational career. For giving me the opportunity to study in Australia, for providing for me, and for being supportive of anything I pursue.

To my sister *Maryam*, thank you for being more excited about my graduation than me and offering support where you could. You are a special person in my heart.

To my best friend *Zoe Morgan*, it is so fantastic to catch up with you! Although we may not see each other as often as we'd like. I feel so lucky to have you by my side.

To *Nor & Colin Lawrence*, thank you for touching my life in ways that you may never know. You two are precious gift sent from God and I am so glad to have met you in my life.

Thesis Abstract

Alzheimer's disease (AD) and Parkinson's disease (PD) are the foremost neurodegenerative diseases (NDs) globally. Accumulated neuropathological insights, supplemented by advanced biomedical tools, underscore intricate, interconnected mechanisms characterising these NDs. A crucial objective is identifying non-invasive biomarkers, ideally from blood, that can detect ND-related pathways before the onset of symptoms. Such biomarkers should be sensitive to early neurological changes and discern the disease state from normal aging processes.

Recent advances in omics techniques, including genomics and transcriptomics, have enriched our understanding of the intricate molecular mechanisms underpinning NDs. These methods have been instrumental in distinguishing subgroups of ND individuals based on specific molecular characteristics, thereby refining biomarker interpretation.

In this study, RNA sequencing (RNA-seq) data was utilised to examine the whole transcriptome from the dorsolateral prefrontal cortex (Brodmann Area 9, or BA9), whole blood (WB), and peripheral blood mononuclear cells (PBMC). The RNA-seq data was sourced from the Sequence Read Archive (SRA), an open-source database, eliminating the need for in-house sequencing. Coupled with single nucleotide polymorphisms (SNPs) from the same samples, I sought to uncover shared pathological events between AD and PD and determine if the molecular pathways discerned in the ND brain are also indicative in blood.

Utilising differential gene expression paired with a random forest-based machine-learning algorithm, our findings indicate limited overlap between AD and PD. Specific AD markers emerged, such as myocyte-specific enhancer factor 2C (Mef2C) in the AD-DLPFC/BA9 and transmembrane P24 trafficking protein 7 (TMED7). Notably, AD blood markers highlighted cilia-mediated TLR4/NF- κ B activation and Fyn. For PD, SNARE complex proteins, like VAMP2 and

STX1A, were predominant, suggesting a linkage between α -syn and SNARE protein redistribution, ultimately contributing to synaptic failure.

Subsequent genome-wide association studies (GWAS) on the samples corroborated some findings, with overlaps noted in the brain and blood for both NDs. Specifically, vascular endothelial growth factor A (VEGF-A) and inositol polyphosphates (InsPs) were identified. The AD blood samples were particularly marked by brain-derived factors known to be associated with A β aggregation and tau pathology.

Further analysis through expression and splicing quantitative trait loci (eQTL and sQTL) mapping elucidated roles of certain GWAS-linked markers in AD and PD risk. Overlaps between GWAS and *cis*-eQTL/*cis*-sQTL indicate shifts in gene expression patterns and alternative splicing events that modulate disease susceptibility in both NDs.

In conclusion, our study underscores the potential of leveraging GWAS and QTL signals to gain insights into the intricate interplay underlying complex diseases like AD and PD. It also emphasizes the importance of further investigations to identify potential therapeutic targets for these predominant NDs.

Abbreviations

A β : Amyloid- β
ABCB1: ATP binding cassette subfamily B member 1
ACE: Angiotensin-converting enzyme
AD: Alzheimer's disease
ADAM10/17/19: ADAM metallopeptidase domain 10/17/19
ADGRL3: Adhesion G protein-coupled receptor L3
ARHGEF2/11: Rho/Rac guanine nucleotide exchange factor 2/11
AJUBA: Ajuba LIM protein
ALP: Autophagy-lysosomal pathway
AMPK: AMP-dependent protein kinase
ANKRD17: Ankyrin repeat domain 17
APOE: Apolipoprotein E
ApoER2: ApoE receptor 2
APP: amyloid precursor protein
AQP: Aquaporin
Arl1: ADP ribosylation factor like GTPase 1
ARL13B: ADP ribosylation factor like GTPase 13B
AS: Alternative Splicing
ATG: Autophagy-related protein
B2M: beta-2-microglobulin
B4GALNT1: Beta-1,4-N-acetyl-galactosaminyltransferase 1
BA9: Brodmann Area 9
BAG3: BAG cochaperone 3
BACE1: β -Site APP cleaving enzyme 1
BCAP31: B Cell receptor associated protein 31
BDNF: brain-derived neurotrophic factor
CAMK2B: Calcium/Calmodulin dependent protein kinase II beta
CARD19: Caspase recruitment domain family member 19
CCSER1: Coiled-coil serine rich protein 1
cDNA: Complementary deoxyribonucleic acid
CHCHD3: Coiled-coil-helix-coiled-coil-helix domain containing 3
CHORDC1: Cysteine and histidine rich Domain containing 1
CMA: Chaperon-mediated autophagy
CNS: Central nervous system
CPLX2: Complexin 2
CREB: Cyclic adenosine monophosphate response element-binding protein

CRIM1: Cysteine rich transmembrane BMP regulator 1
 CSMD1: CUB and sushi multiple domains 1
 CXCL12: C-X-C motif chemokine ligand 12
 DAMPs: Danger-associated molecular patterns
 DGE: Differential gene expression
 DLP1: Dynamin like protein 1
 DLPFC: Dorsolateral prefrontal cortex
 DNAJB1: DnaJ heat shock protein family (Hsp40) member B1
 DNM1/2: Dynamin1/2
 ECE: Endothelin converting enzyme
 EEF2: eukaryotic translation elongation factor 2
 EIF3E: Eukaryotic translation initiation factor 3 subunit E
 eMI: Endosomal micro-autophagy
 EOAD: Early-onset autosomal AD
 EPHA6: EPH receptor A6
 ER: endoplasmic reticulum
 ERGIC: ER-Golgi intermediate compartment
 ESE: Exonic splicing enhancers
 ESS: Exonic splicing silencer
 EVs: Extracellular vesicles
 FAAH: fatty acid amide hydrolase
 FCγR2A/B: Fc fragment of IgG receptor IIa/b
 FRMD4A: FERM domain containing 4A
 FZD3: Frizzled class receptor 3
 GABRB1: Gamma-aminobutyric acid type A receptor beta1 subunit
 GATK: genome analysis toolkit
 GBF1: Golgi brefeldin a resistant guanine nucleotide exchange factor 1
 GFAP: Glial fibrillary acidic protein
 GI: Genetic interaction
 GPM6A: glycoprotein M6A
 GRB2: Growth factor receptor bound protein 2
 GRIK4: Glutamate ionotropic receptor kainate type subunit 4
 GRIN2A/B/D: Glutamate ionotropic receptor NMDA type subunit 2A/B/D
 GSH: glutathione
 GSK3β: Glycogen synthase kinase-3 beta
 GWAS: Genome-wide association studies
 HLA-A/B/C/E/G: Major histocompatibility complex, class I, A/B/C/E/G

HLA-DRA: Major histocompatibility complex, class II, DR alpha
 HLA-DRB1/5: Major histocompatibility complex Class II, DR beta 1/5
 HLA-DPA1: Major histocompatibility complex, class II, DP alpha 1
 HLA-DQA1: Major histocompatibility complex, Class II, DQ alpha 1
 HLA-DQB1: Major histocompatibility complex, Class II, DQ beta 1
 HNRNPM: Heterogeneous nuclear ribonucleoprotein M
 HNRNPUL1: Heterogeneous nuclear ribonucleoprotein U-like protein 1
 HOMER1/2/3: homer scaffold protein 1/2/3
 HSP90AA1: Heat shock protein 90 alpha family class A member 1
 HSPA1A/B: Heat shock protein family A (Hsp70) member 1A/B
 HSPA8: Heat shock protein family A (Hsp70) member 8
 HSPGs: Heparan sulfate proteoglycans
 IDE: Insulin-degrading enzyme
 ILs: Interleukins
 INPP5B/D: inositol polyphosphate-5-phosphatase B/D
 IP3Rs: Inositol triphosphate receptors
 ISE: Intronic splicing enhancer
 ISS: Intronic splicing silencer
 KEGG: Kyoto Encyclopedia of Genes and Genome
 KIF1A: Kinesin family member 1A
 KIR3DS1: Ig domains and short cytoplasmic tail 1
 LAG3: lymphocyte activation gene 3
 Lamp2A: Lysosome-associated membrane protein type 2A
 LDLR: Low-density lipoprotein receptors
 LIMK1: LIM domain kinase 1
 LOAD: Late-onset AD
 LRRK2: Leucine-rich repeat kinase 2
 MAF: Minor allele frequency
 MAMs: Mitochondria-associated membranes
 MAPK: p38 mitogen-activated protein kinase
 MAPKAPK2: MAPK activated protein kinase 2
 MAPT: microtubule-associate protein tau
 MEF2C: Myocyte enhancer factor 2
 mEOD: Mendelian EOD
 miR-rSNP: MicroRNA regulatory SNP
 ML: Machine learning
 MMP: Matrix metalloproteinase

MOBP: Myelin associated oligodendrocyte basic protein mRNA: Messenger RNA
 MTBD: Microtubule-binding domain
 MTFR1: mitochondrial fission regulator 1
 mTORC1: Mammalian target of rapamycin complex 1
 MYO19: myosin XIX
 MZS: Maximum Z score
 NBEAL1: Neurobeachin like 1
 ND: Neurodegenerative disease
 NDUFC2: NADH dehydrogenase [ubiquinone] 1 subunit C2
 NDUFS3: NADH dehydrogenase [ubiquinone] iron-sulphur protein 3
 NEP: Neprilysin
 NFATC2: Nuclear factor of activated T cells 2
 NF- κ B: Nuclear factor- κ B
 NFTs: Neurofibrillary tangles
 NGS: Next-generation sequencing
 NHLRC2: NHL repeat containing 2
 NLRP3: NOD, LRR- and pyrin domain-containing 3
 nmEOAD: Nonmendelian EOAD
 NPM1: Nucleophosmin 1
 PAK1: P21 (RAC1) activated kinase 1
 PARK7: Parkinson's disease protein 7
 PBMC: Peripheral blood mononuclear cells
 PCR: Polymerase chain reaction
 PD: Parkinson's disease
 PHFs: Paired helical filaments
 PIK3CD: Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta
 pre-mRNA: Precursor mRNA
 PSEN: Presenilin gene
 p-tau: Hyperphosphorylated microtubule-associated tau protein
 PTMs: Post-translation modifications
 Q-Q: Quantile-quantile
 RAC1: Rac family small GTPase 1
 RANBP1: RAN binding protein 1
 RELA: RELA proto-oncogene, NF- κ B subunit
 RNA: Ribonucleic acid
 RNA-seq: RNA-sequencing
 RNF225: Ring finger protein 225

RNP: Ribonucleoprotein complex
 RNS: Reactive nitrogen species
 ROCK2: Rho-associated coiled-coil containing protein kinase 2
 ROS: Reactive oxygen species
 RPS5: Ribosomal protein S5
 rSNP: Regulatory SNP
 SF3A2: Splicing factor 3a subunit 2
 SF3B4: Splicing factor 3b subunit 4
 SFPQ: Splicing factor proline and glutamine rich
 SHNAK3: SH3 and multiple ankyrin repeat domains 3
 SKP1: S-phase kinase associated protein 1
 SLC1A2: Solute carrier family 1 member 2
 SNAP23/25/29/47: Synaptosome associated protein 23/25/29/47
 SNAPIN: SNAP associated protein
 SNARE: N-ethylmaleimide-sensitive factor attachment protein receptor
 SNCA: α -synuclein gene
 SNP: Single nucleotide polymorphism
 snRNP: Small nuclear ribonucleoprotein
 SORL1: Sortilin-related receptor 1
 SRA: Sequence Read Archive
 SRSF2/4/7/9: Serine and arginine rich splicing factor 2/4/7/9
 srSNP: Structural RNA SNP
 STX1A/B: syntaxin 1A/B
 SYT1/2: synaptotagmin 1/2
 TGN: Golgi/trans-Golgi network
 TFEB: Transcription factor EB
 TLR: Toll-like receptors
 TMED7: Transmembrane P24 trafficking protein 7
 TMM: Trimmed Mean of M-values
 TNF- α : Tumour necrosis factor-alpha
 TNT: Tunnelling nanotube
 TOM20: Outer membrane 20
 TRPML: Transient receptor potential channel mucolipin
 TUBA1A/B: Tubulin alpha 1A/B
 ULK1: Unc-51-like kinase 1
 UPS: Ubiquitin-proteasome system
 UTR: Untranslated region

VAMP1/2/3: Vesicle-associated membrane protein 1/2/3
VDAC1: Voltage-dependent anion channel 1
VEGF-A/B: Vascular endothelial growth factor A/B
VEP: Variant Effect Predictor
VLDLR: Very low-density lipoprotein receptor
VTI1B: t-SNAREs 1B
WASF1/3: WAS protein family member 1/3
WB: Whole blood
WDL: workflow description language
WNT5A: Wnt family member 5A
ZAP70: Zeta chain of T cell receptor associated protein kinase 70
ZC3H4: Zinc finger CCCH-type containing 4

Table of Contents

Chapter 1 Background	1
1.1 Neurodegenerative disease (ND)	1
1.2 Signature discovery techniques for NDs	2
1.2.1 RNA-sequencing technology	3
1.2.1.1 Differential gene expression (DGE)	3
1.2.1.2 Machine Learning (ML)	5
1.2.1.3 Alternative splicing	6
1.2.1.4 Single nucleotide variants	9
1.2.2 Tissue resources for transcriptome biomarker discovery	11
1.3 Alzheimer's Disease (AD)	13
1.3.1 Propagation of A β proteinopathy in LOAD	15
1.3.1.1 APP structure and endo-lysosomal trafficking	16
1.3.1.2 α -secretases endo-lysosomal trafficking	17
1.3.1.3 β -secretase endo-lysosomal trafficking	18
1.3.1.4 γ -secretase endo-lysosomal trafficking	19
1.3.1.5 APP proteolytic processing and A β production	20
1.3.2 A β function under homeostatic and pathological conditions	21
1.3.2.1 A β deposits and induced-tau pathology	23
1.3.3 Clearance or removal of A β deposits	27
1.3.3.1 A β receptor-mediated clearance	27
1.3.3.2 A β clearance mediated by proteases	28
1.3.3.3 A β clearance mediated by autophagy-lysosome system	29
1.3.3.4 A β clearance mediated by neuroglia activation	30
1.3.3.5 A β clearance mediated by cerebrospinal fluid (CSF)	31
1.4 Parkinson's Disease (PD)	32
1.4.1 Propagation of α -synucleinopathy in PD	33
1.4.1.1 α -syn structure	33
1.4.1.2 α -syn structural plasticity and interactions	34
1.4.1.3 Transmission of toxic forms of α -syn in the brain	36
1.4.2 Misfolding and aggregation of α -syn: implications in PD pathogenesis	37
1.4.2.1 α -syn aggregation and autophagy-lysosomal pathway impairment	38
1.4.2.2 α -syn aggregation and glia-mediated neuroinflammation	38
1.4.2.3 α -syn aggregation and mitochondria dysfunction	39
1.4.2.4 α -syn aggregation and synaptic dysfunction	40
1.4.2.5 α -syn aggregation alters modulation of P-body components	40
1.5 Summary	42
Chapter 2 Overlap between central and peripheral transcriptomes in Parkinson's disease but not Alzheimer's disease	44
2.1 Introduction	45
2.2 Materials and methods	47
2.2.1 Data collection	47
2.2.2 Workflow implementation and data processing	47
2.2.3 Transcriptomic data normalisation and expression analysis	48
2.2.4 Analysis of genes associated with NDs by machine learning (ML)	49
2.2.5 Statistical assessment of overlapping genes	50
2.2.6 Gene ontology and pathway enrichment analysis	50
2.2.7 Gene-gene interaction and gene co-expression network analysis	51
2.3 Results	51

2.3.1 Non-overlapping sample subset selection using differentially expressed genes-----	51
2.3.2 ML-based ranking of genes in BA9 (or DLPFC) from AD and PD-----	59
2.3.3 Diagnosis of AD and PD by profiling peripheral blood biomarkers using ML-----	61
2.3.4 Correspondence between peripheral blood and brain in NDs-----	65
2.4 Discussions -----	67
2.5 Conclusions-----	72
2.6 Potential limitations and future prospective -----	72
Chapter 3 Transcriptome based genetic variant discovery from peripheral and central tissues to characterise Alzheimer's and Parkinson's disease -----	75
3.1 Introduction -----	75
3.2 Materials and methods-----	76
3.2.1 Data collection-----	76
3.2.2 RNA-seq mapping and data processing-----	77
3.2.3 RNA-based variant detection using GATK pipeline -----	77
3.2.4 Annotation and distribution of variants across the genome -----	78
3.2.5 Quality control (QC) and genome-wide association studies (GWAS)-----	79
3.2.6 Locus gene-based pathway analysis -----	79
3.3 Results-----	80
3.3.1 Genomic variant detection from RNA-seq data workflow -----	80
3.3.2 Replication studies for SNPs detected by RNA-seq with those detected by genome sequencing (not same individuals) -----	82
3.3.3 Genomic location of SNPs detected by RNA-seq for each comparison -----	83
3.3.4 Genome-wide association studies (GWAS) in BA9 (or DLPFC) from AD and PD-----	86
3.3.5 Genomic variants in peripheral blood associated with AD and PD -----	90
3.3.6 Functional pathways associated with AD and PD in BA9 (or DLPFC) -----	93
3.3.7 Functional pathways influencing AD and PD in peripheral blood -----	98
3.4 Discussions -----	106
3.5 Conclusions-----	113
3.6 Potential limitations and future prospective -----	115
Chapter 4 Exploring the effect of genetic variants (SNPs) on alternative splicing and gene expression patterns in blood and brain of Alzheimer's and Parkinson's disease -----	117
4.1 Introduction -----	117
4.2 Materials and methods-----	119
4.2.1 RNA-based SNP calling and quality control (QC) -----	119
4.2.2 Gene expression (GE) analysis-----	119
4.2.3 Prediction of alternative splicing (AS) features-----	120
4.2.4 Expression quantitative trait loci (eQTL) mapping-----	120
4.2.5 Identification of splicing quantitative trait loci (sQTL)-----	121
4.2.6 Gene enrichment and interaction analysis -----	121
4.3 Results-----	121
4.3.1 Identification of <i>cis</i> -expressed QTLs in central brain and peripheral blood from AD and PD-----	121
4.3.2 Identification of <i>cis</i> -spliced QTLs in central brain and peripheral blood from AD and PD -----	130
4.3.3 Comparative analysis of <i>cis</i> -eQTLs and <i>cis</i> -sQTLs identified via RNA-seq and genome sequencing from distinct cohorts -----	139
4.4 Discussions -----	140
4.5 Potential imitations and future prospective-----	143

List of Figures

Figure 1.1 RNA-sequencing (RNA-seq): gene expression and feature selection analysis.	4
Figure 1.2 Schematic figure representing five modes of alternative splicing of pre-mRNA.	8
Figure 1.3 Genomic location of the SNPs across the population and their detection using RNA-seq data.	10
Figure 1.4 Domain architecture of the amyloid precursor protein (APP695).	16
Figure 1.5 APP processing via the non-amyloidogenic and amyloidogenic pathways.	20
Figure 1.6 Structure of the full-length tau protein and molecular mechanisms for cell-to-cell transmission of pathogenic tau.	25
Figure 1.7 Pathways associated with A β clearance in the brain.	27
Figure 1.8 The structure of monomeric α -synuclein (α -syn).	34
Figure 1.9 α -syn aggregation affecting critical cellular functions in the brain.	37
Figure 2.1 Schematic representation showing all steps necessary for RNA-seq analysis.	53
Figure 2.2 Score plots for the first two Principal Components (PCs) of comparing groups.	54
Figure 2.3 Hierarchical clustering for different tissue in comparing groups.	55
Figure 2.4 Schematic representation of ANOVA test for comparing groups.	56
Figure 2.5 Schematic representation for DLPFC-BA9 tissue in comparing groups.	60
Figure 2.6 Schematic representation for PBMC tissue in comparing groups.	63
Figure 2.7 Schematic representation for WB tissue in comparing groups.	64
Figure 2.8 Venn diagrams illustrate the overlapping features.	66
Figure 2.9 Overlapped biological processes (BP) of Gene Ontology terms.	67
Figure 3.1 Schematic representation showing all steps necessary for RNA-seq based genomic variant discovery and downstream analysis.	81
Figure 3.2 Repartition of the SNPs detected by RNA-seq among different genomic regions.	83
Figure 3.3 Total number of SNPs detected by RNA-seq among comparing groups.	84
Figure 3.4 Quantile-Quantile plots (QQ-Plots) of comparing genome-wide association study (GWAS) groups.	86
Figure 3.5 Density plots of single nucleotide polymorphism (SNP) distribution on 22 chromosomes.	88
Figure 3.6 Manhattan plots summarizing the GWAS results for comparing groups.	91
Figure 3.7 Venn diagram illustrate the overlapping single nucleotide polymorphisms (SNPs) among study groups.	94
Figure 3.8 Kyoto encyclopedia of genes of genomes (KEGG) enrichment analysis for comparing groups.	97
Figure 3.9 Shared genetic variants (SNPs) between comparing groups.	101
Figure 3.10 Significant overlapped kyoto encyclopedia of genes of genomes (KEGG) enriched pathways between study groups.	105
Figure 4.1 Quantile-Quantile plots (QQ-Plots) of local and distal eQTLs in AD and PD cohorts.	123
Figure 4.2 Waterfall plot of the top 30 upregulated or downregulated <i>cis</i> -eQTLs in DLPFC/BA9 samples of AD and PD cohorts.	125
Figure 4.3 Waterfall plot of the top 30 upregulated or downregulated <i>cis</i> -eQTLs in PBMC samples of AD and PD cohorts.	127
Figure 4.4 Waterfall plot of the top 30 upregulated or downregulated <i>cis</i> -eQTLs in WB samples of AD and PD cohorts.	129
Figure 4.5 Quantile-Quantile plots (QQ-Plots) of local and distal sQTLs in AD and PD cohorts.	131
Figure 4.6 Frequency of AS events in <i>cis</i> -sQTLs identified for DLPFC/BA9 samples of AD and PD cohorts.	131
Figure 4.7 Volcano plot showing the difference of <i>cis</i> -sQTLs related AS-events in DLPFC/BA9 samples of AD and PD cohorts.	132
Figure 4.8 Frequency of AS events in <i>cis</i> -sQTLs identified for PBMC samples of AD and PD cohorts.	135
Figure 4.9 Volcano plot showing the difference of <i>cis</i> -sQTLs related AS-events in PBMC samples of AD and PD cohorts.	135
Figure 4.10 Frequency of AS events in <i>cis</i> -sQTLs identified for WB samples of AD and PD cohorts.	138
Figure 4.11 Volcano plot showing the difference of <i>cis</i> -sQTLs related AS-events in WB samples of AD and PD cohorts.	139

List of Tables

Table 2.1 Number of tissue-specific samples per phenotype.	51
Table 2.2 Top ten differentially expressed genes (DEGs) for comparing groups.	57
Table 2.3 Number of differentially expressed genes (DEGs) and enriched GO-BPs from a selected subset of samples.	58
Table 3.1 Top ten single SNP results from a genome-wide scan in terms of a smaller number of clumps of correlated SNPs per tissue-specific samples in case-control/case-case studies. Genes information related to each SNP were taken from dbSNP (https://www.ncbi.nlm.nih.gov/snp/).	89
Table 3.2 <i>P</i> -values of associated genes in comparing groups per tissue.	95
Table 3.3 <i>P</i> -values for overlapping risk loci among AD and PD in central brain and peripheral blood.	99
Table 3.4 <i>P</i> -values of overlapped genes associated with AD and PD in central brain and peripheral blood.	103
Table 4.1 Number of SNP genotypes, <i>cis</i> - and <i>trans</i> -eQTLs, and sQTLs for each group and tissue.	121
Table 4.2 <i>P</i> -values of <i>cis</i> -eQTL for locus identified for each contrast per tissue.	126
Table 4.3 <i>P</i> -values of overlapped <i>cis</i> -eQTLs between AD and PD in central brain and peripheral blood.	127
Table 4.4 <i>P</i> -values of <i>cis</i> -sQTL for locus identified for each contrast per tissue.	133
Table 4.5 <i>P</i> -values of overlapped <i>cis</i> -sQTLs between AD and PD in central brain and peripheral blood.	136

List of Appendices

Appendix 1 Table contains information of samples downloaded from National Centre of Biotechnological Information's Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra/).	179
Appendix 2 Table provides a description of the selected SRA samples through samples subtyping analysis.	200
Appendix 3 Table provides information of RNA-seq samples with ~150bp read lengths downloaded from SRA.	214

Chapter 1 Background

1.1 Neurodegenerative disease (ND)

Neurodegenerative disease (ND) is an umbrella term for a range of conditions primarily characterised by the progressive loss of structure and function of neurons in the central and peripheral nervous system [1]. ND is a common cause of morbidity and cognitive impairment in older adults worldwide, which is characterised by extra- or intra-cellular deposits of disease-specific misfolded protein in the central nervous system (CNS) [2]. The primary process of protein misfolding and aggregation involves rearranging the structure of the protein into a series of β -strands stabilised by hydrogen bonding and hydrophobic interactions [3]. Then each molecule can adopt many varied dysfunctional structures (soluble non-native oligomers and large insoluble protein deposits (fibrils and plaques)) which appear to be especially toxic and further reduce the effectiveness of protein folding systems by sequestering chaperones and other factors, giving rise to various type of NDs [4].

Alzheimer's disease (AD) and Parkinson's disease (PD) are well-known age-related NDs that are characterised by misfolded protein deposits [5]. These misfolded proteins appear to spread prion-like, with progressive neuronal loss in specific brain areas [6]. The diagnosis of AD and PD is based on clinical criteria that utilise a structured scoring system for examining disease-related symptoms. However, clinical diagnostic criteria cannot detect NDs in the early or preclinical phases when prospective disease-modifying treatments are potentially more effective. In particular, the large variety of comorbidities in age-related NDs has made early disease diagnosis, identification of individuals at risk, and disease progression monitoring challenging. Therefore, investigation of comparative signatures in affected post-mortem brain versus peripheral tissues (whole blood (WB), and peripheral blood mononuclear cells (PBMC)) from live donors

(depending on the disease stage), may allow researchers to report stages/phases of proteinopathies and their dependent pathogenic pathways that occur before the onset and during the progression of NDs [7].

1.2 Signature discovery techniques for NDs

Early detection of the onset of neurodegeneration is vital as it can provide a chance for an early treatment that may help prevent the further progression of AD and PD [8, 9]. Molecular diagnosis has emerged as a powerful technique that can be helpful for the early detection of the onset of NDs. One of the pivotal approaches in molecular diagnostics involves using biomarkers to detect or predict the presence or onset of a specific disorder [9]. These biomarkers, once identified, can potentially guide therapeutic interventions or highlight new therapeutic targets. With the advent of advanced technologies, “-omics” studies are becoming instrumental in this endeavour. “-omics” or high-throughput sequencing techniques, including but not limited to next-generation sequencing, are now extensively employed to elucidate molecular pathogenesis at the genomic, transcriptomic, proteomic, metabolomic, and phenomic levels [10]. Unlike the genome, which remains relatively fixed throughout the lifetime of an organism, the transcriptome, a set of ribonucleic acid (RNA) molecules in a cell, can readily change depending on certain environmental and internal factors, such as nutritional status, hormonal signals, pathogens, toxins and drugs, oxidative stress, pH and ionic strength, light and circadian rhythms, developmental signals, cell-to-cell communication, tissue injury and inflammation, and aging. Thus, studies of the transcriptome through RNA-sequencing technology provide insight into what is occurring in a specific tissue/ or cell at a particular moment.

1.2.1 RNA-sequencing technology

1.2.1.1 Differential gene expression (DGE)

RNA-sequencing (RNA-seq) is often used to profile differential gene expression (DGE) in a case-control paradigm. DEG are genes whose transcripts are significantly different across two or more conditions [11]. The standard workflow of a DGE assay begins in the laboratory, RNA extraction from the sample. Once isolated, this total RNA undergoes either mRNA enrichment to concentrate the messenger RNA or ribosomal RNA (rRNA) depletion to remove the predominant rRNA. This purified RNA is then reverse transcribed to form complementary DNA (cDNA) in a process termed cDNA synthesis. To ensure that the target cDNA is adequately represented and to obtain sufficient quantities for sequencing, a polymerase chain reaction (PCR) is often employed during library preparation. It's crucial to note that while PCR amplification facilitates adequate sequencing depth, it can also introduce biases that might impact quantification [12]. Subsequently, the cDNA undergoes an adaptation where known sequences, or adaptors, are added, preparing it for high-throughput sequencing. This final, adaptor-ligated product forms a sequencing library. The library is sequenced on a high-throughput platform, generally targeting a read depth of 10-30 million reads per sample (Figure 1.1 A) [11, 12]. The final steps are computational: aligning and/or assembling the sequencing reads to a reference genome (transcriptome), quantifying reads that overlap transcripts, filtering, and normalising between samples, and using DEG tools to identify significant changes in the expression levels of individual transcripts between sample groups (Figure 1.1 B) [12].

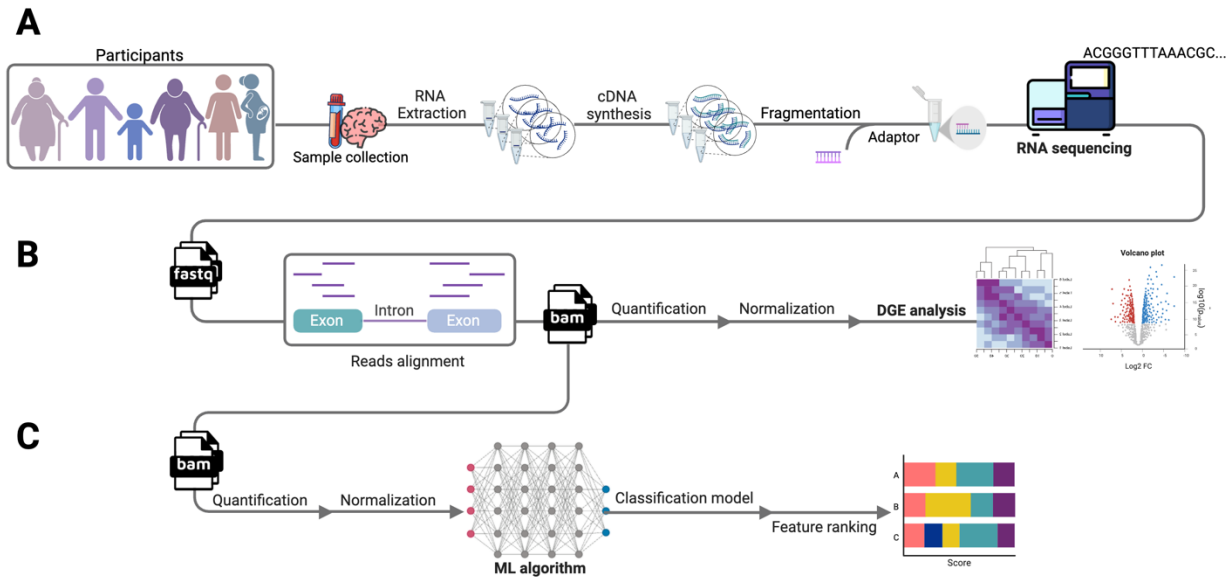


Figure 1.1 RNA-sequencing (RNA-seq): gene expression and feature selection analysis.

A) An overview of the main steps for RNA-sequencing (RNA-seq). Once the RNA sample is obtained, the first step in the technique involves converting the population of RNA into complementary DNA (cDNA). This is done by reverse transcription, allowing the RNA into the sequencing workflow. The cDNA is then fragmented (a cDNA library), and adapters are added to each end of the fragments. These adapters contain functional elements which permit sequencing. Following amplification, size selection, clean-up and quality-checking processes, the cDNA library is then analysed by omics or high throughput sequencing technologies, producing short sequences that correspond to all or part of the fragment from which it was derived. **(B)** Computational analysis for differential gene expression (DGE) begins by aligning raw sequence data in FASTQ format to the reference sequence, with results stored in the binary alignment/map (BAM) file format. Quantification tools then assign the reads to features. Normalisation, usually methods being used within the quantification or expression modelling tools, adjusts raw data to account for factors that can skew direct comparisons of expression measures. These factors include gene length, as longer genes generally have more reads mapped than shorter ones; GC-content, since variations can affect sequencing efficiency; and sequencing depth, where a sample with more total reads might appear to have higher expression even when normalised. After considering these factors, gene expression is accurately modelled. A list of differentially expressed genes or transcripts is generated for further visualisation and interpretation. **(C)** Feature selection, one of the main components of feature engineering, is selecting the most important features to input in machine learning algorithms.

Feature selection techniques are employed to reduce the number of input variables by eliminating redundant or irrelevant features and narrowing down the set of features to those most relevant to the machine learning model.

By comparing the mean expression values of different groups, RNA-seq DEG tools can measure the magnitude of differential expression between two or more conditions based on read counts from replicated samples, i.e., calculate the fold change (FC) of read counts, and considering the differences in sequencing depth and variability [13]. This is followed by a statistical test based on the null hypothesis that the difference is close to zero, which would mean that there is no difference in the gene expression values that the conditions could explain. A p -value, the probability of obtaining an experimental result at least as extreme as the one observed under the null hypothesis, would then be obtained [13]. Therefore, RNA-seq DEG tools determine genes that have statistically significant differences while also providing information related to the expression level and pairwise magnitude of difference for each gene [14]. However, one of the biggest problems in RNA-seq DGE analysis is that all the DGE analysis methods invite almost all genes into DGE analysis without conducting feature selection for high-dimensional RNA-seq data. This problem is known as the “curse of dimensionality” and substantially impacts analysis accuracy and recall due to overfitting that can mistake small fluctuations for significant variance in the data and lead to spurious findings [15].

1.2.1.2 Machine Learning (ML)

Machine Learning (ML), a branch of artificial intelligence, leverages algorithms and statistical models to enable computers to execute tasks without explicit commands. In doing so, ML can discern patterns, forecast outcomes, and enhance data interpretation by processing vast datasets, with its applications spanning areas like disease diagnosis and health and safety [16], as depicted in Figure 1.1 C. This capability becomes indispensable in the context of RNA sequencing (RNA-

seq). Given the voluminous data generated through RNA-seq, ML aids in distilling key insights. One cornerstone of integrating ML with RNA-seq is 'feature engineering,' encompassing both feature extraction and selection. In this context, a "feature" signifies an individual measurable trait, such as gene transcripts, pivotal for understanding disease mechanisms [17]. Feature extraction crafts new variables from raw data informed by domain-specific insights, while feature selection zeroes in on the most telling features to craft efficient models [18].

In investigating neurodegenerative diseases (NDs), ML-driven feature selection stands out as a compelling technique for parsing expansive datasets, playing a crucial role in diagnosis, prognosis, and therapeutic innovations [17]. Notably, ML techniques eclipse traditional methods due to their adeptness at managing data marked by high non-linear correlations, resilience against noise, ease of calibration, and consistency in results [18, 19]. Utilized to cherry-pick a subset of genes symbolizing inherent dataset structures, ML-enhanced feature selection optimizes the signal-to-noise ratio, leading to more accurate clustering than the traditional differential gene expression (DEG) analysis, which generally pinpoints genes with significant shifts in expression under varied conditions [19].

Beyond its utility in DEG analysis and ML-based feature selection, RNA-seq data offers a broader spectrum of applications. For instance, it serves as a pivotal resource for investigating mRNA splicing events, which can elucidate the complexity of gene regulation. Additionally, RNA-seq aids in detecting single nucleotide variations, providing insights into genetic variability and potential mutations.

1.2.1.3 Alternative splicing

Most protein-coding genes in eukaryotes are transcribed as precursor mRNA (pre-mRNA) that consists of protein-coding regions (exons) with non-coding regions (introns) (Figure 1.2 A) [20].

Alternative splicing is the process by which pre-mRNA exons can be spliced into different arrangements to produce structurally and functionally distinct mRNA and protein variants. Pre-mRNA alternative splicing occurs in a large ribonucleoprotein complex (RNP), a spliceosome. The spliceosome is a dynamic complex mainly consisting of five small nuclear ribonucleoproteins (snRNPs) (U1, U2, U4/U6, U5) and approximately 100 proteins that recognise and assemble on each intron to remove introns from a pre-mRNA [20]. In humans, there are a variety of different ways that alternative splicing events occur. Exon skipping (also known as cassette exons), which is reported to be the most common alternative splicing event in mammalian cells, results in the complete skipping of one or more exons (Figure 1.2 (A1)). Mutually exclusive exons are a rare subtype where two or more splicing events are not independent (Figure 1.2 (A2)). They are executed or disabled in a coordinated manner. Alternative 5' splice sites (alternative donors) are the usage of an alternative 5' donor site, that changes the 3' boundaries of the upstream exon (Figure 1.2 (A3)). Similarly, alternative 3' splice sites (alternative acceptors) lead to a change of the 5' boundaries of the downstream exon (Figure 1.2 (A4)). Intron retention (IR) (Figure 1.2 (A5)) is the process where one or more introns remain un-spliced in the mRNA [21].

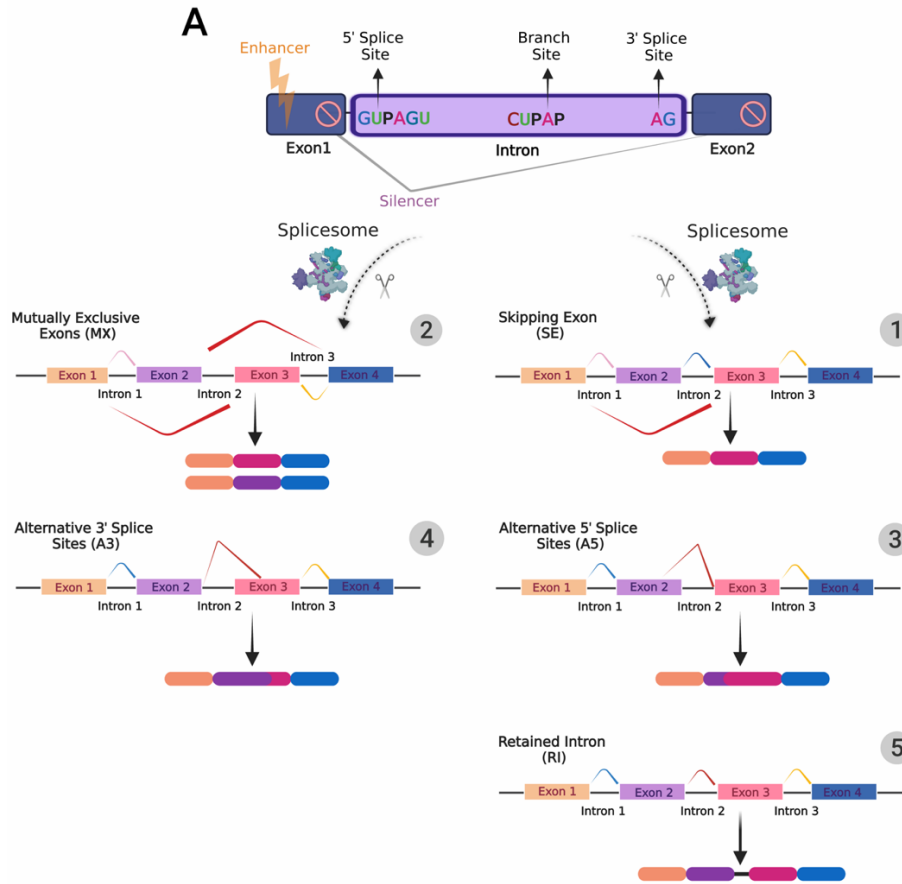


Figure 1.2 Schematic figure representing five modes of alternative splicing of pre-mRNA.

(A) In the pre-mRNA, several conserved motifs in the nucleotide sequences near the intron-exon boundaries act as essential splicing signals. These signals- *GU* and *AG* dinucleotides at the 5' ends of the introns are called splice donor sites or 5' splice sites, and at the 3' ends of the introns are called splice acceptor sites or 3' splice sites, followed by a polypyrimidine tract (*Py*)_n and an *A* nucleotide at the branch site; are involved in the excision of introns from pre-mRNA and in the joining of the exons. Additional regulatory elements known as 'enhancers' and 'silencers' are also required to allow normal splicing of exonic sequences. Next, five types of alternative splicing events, including (1) skipping exon, (2) mutually exclusive exons, (3) alternative 5' splice site, (4) alternative 3' splice site, and (5) retained intron, are shown. In these graphics, exons are represented by boxes and introns by lines. For each category, the different splicing reactions are symbolised by a red arrow.

These alternative splicing events are regulated by interactions between cis-acting and trans-acting elements that either promote (enhancers) or inhibit (silencers) splicing activity of nearby splice

sites [21]. These elements include splicing enhancers/ silencers that are located either in exons (exonic splicing enhancers (ESEs)/ ESSs) or in introns (intronic splicing enhancers (ISEs)/ ISSs) [20, 21]. Notably, mutations that lead to disruptions of either *trans*-acting regulatory proteins or *cis*-acting regulatory sequences could lead to aberrant splicing and noticeable changes in the transcriptome profile of a specific tissue. In addition, both the ageing process and/or environmental factors could influence alternative splicing events and contribute to the onset of disease [22]. Therefore, analysis of alternative splicing in NDs will improve understanding the transcriptome changes with cognitive ageing and its related disorders. Computational analysis of RNA-seq-based alternative splicing consists of three main steps: detection, statistical comparison, and effect prediction. Software packages for detecting splicing alterations are broadly broken down into those that only identify events found in annotated transcripts and those that additionally detect novel splice events [23]. These packages also include functions for testing differential splicing (DS) events. DS comprises the analysis of alternative splicing in which event or isoform abundance is compared across two or more conditions [23].

1.2.1.4 Single nucleotide variants

Age-associated transcriptome profile changes can also be influenced by genetic variants. Single nucleotide polymorphisms (SNPs) are variations of a single nucleotide base (Figure 1.3 A) that arise from either transitions (C/T or G/A) or transversions (C/G, C/A, T/A, or T/G) at the same position between individual genomic DNA sequences [24]. As the predominant type of DNA polymorphism, SNPs can be found throughout the genome, including intergenic region (regions between genes), coding sequences (exons), or non-coding regions (introns, 5'UTR, 3'UTR, or exon-intron splicing sites) [24,25]. Depending on their genomic location (Figure 1.3 A), SNPs have the potential to alter all steps of gene expression. When present within transcriptional regulatory

elements, they can affect mRNA expression. They can also impact mRNA splicing, nucleocytoplasmic export, stability, and translation. SNPs located in non-coding regions of the genome may affect transcription factor (TF) binding affinity or disrupt the regulatory elements accessibility and function [26]. When present within a coding sequence and leading to an amino acid change (referred to as a non-synonymous SNP or mutation), they can modify the protein's activity [25]. Hence, due to their high frequency and binary variation patterns (two possible variations (alleles) for a specific position), SNPs have been widely used as generic markers to predict an individual's response to certain drugs, susceptibility to environmental factors, and risk of developing numerous complex diseases, including NDs [26].

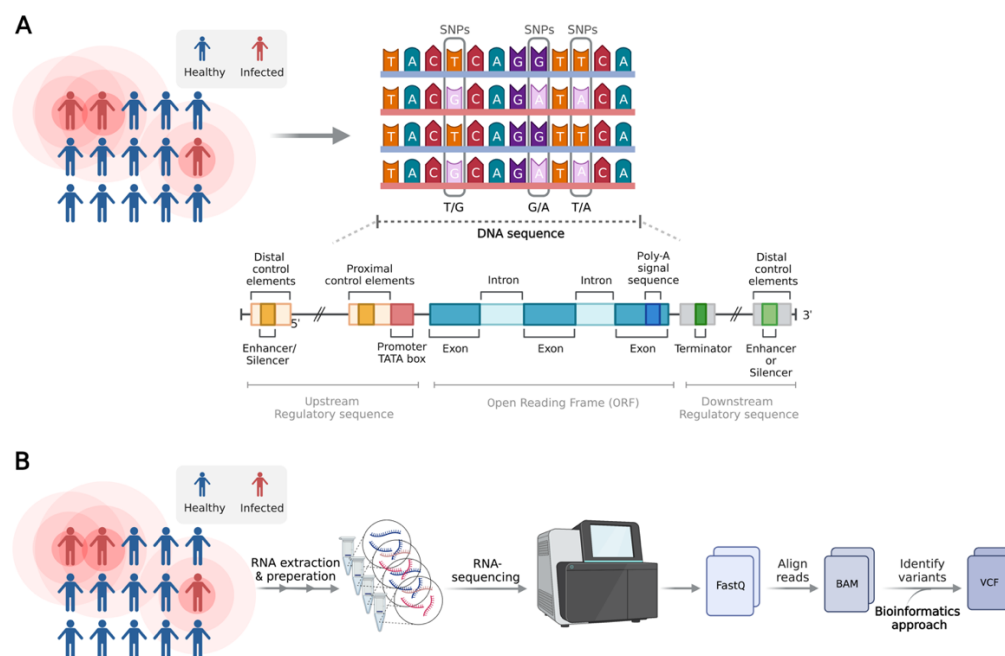


Figure 1.3 Genomic location of the SNPs across the population and their detection using RNA-seq data.

(A) Single nucleotide polymorphisms (SNPs), marked with pink colour, are the most common variation in the human genome that gives rise to different alleles containing alternative bases at a given nucleotide within a locus. SNPs can be found in the promoters of protein-coding and non-coding RNA known as regulatory SNPs (rSNPs) and microRNA regulatory SNPs (miR-rSNP), respectively, that affect gene expression. In addition, SNPs located within the coding region of the precursor mRNAs (exons and introns), mature mRNA (5' untranslated region [UTR], coding sequence,

and 3' UTR), and primary, precursor, and mature miRNAs are termed structural RNA SNPs (srSNPs) and miR-srSNPs, respectively. The srSNPs affect alternative splicing, translation, stability, amino acid sequence, structure, and function of proteins and the interaction between mRNA/miRNAs. Finally, the miR-srSNPs affect the structure, processing, and interaction between miRNAs/mRNAs. **(B)** Next-generation sequencing provides a powerful way to identify novel SNPs and call known SNPs in transcriptome samples. Accurately calling these SNPs requires high-quality sequencing data, high coverage, and a thorough bioinformatics approach. All variant calling tools can be applied to individual samples after alignment (BAM) and pre-processing are complete. The position of the variants called for each sample is reported in the variant calling format (VCF) files.

SNPs can be identified by whole genome (WG) or whole exome (WE) sequencing [27]. However, in the absence of WGS or WES approaches, SNP calling from RNA-seq (Figure 1.3 B) will offer a suitable alternative to either approach or be used to validate SNPs detected from either WGS or WES [27, 28]. In addition, most SNPs identified via this analysis are in the coding regions of the genome, wherein variants are most likely to result in phenotypic changes and undergo selection pressure [28]. Therefore, SNPs can be accurately discovered both within the same gene and between different samples by aligning typical 150 bp read lengths of RNA-seq to a reference genome and application of an SNP caller tool [27, 28]. The SNP caller selects the reads that match any of the reference sequences. Reads with no match or matching ribosomal DNA (for RNA sequencing) are not used by the SNP caller.

1.2.2 Tissue resources for transcriptome biomarker discovery

Patients with AD and PD have neuropathology with the early loss of neurons in specific brain regions and then progressive loss across regions correlated with increasing severity of cognitive impairment in both diseases. Finding biomarkers associated with mechanisms and pathophysiological changes in particular brain regions would help us to identify patients at high risk for early and rapid cognitive impairments and potential therapeutic pathways.

The dorsolateral prefrontal cortex (DLPFC) is located within a section of the middle frontal gyrus (MFG). It encompasses parts of the frontal cerebral cortex, specifically Brodmann areas 46 (BA46) and 9 (BA9). These Brodmann areas are numbered regions of the cerebral cortex differentiated by their unique cellular organization and histological features. The DLPFC plays a crucial role in executive control functions, including task switching, task-set reconfiguration, interference prevention, inhibition, planning, and working memory [29]. Notably, mild to moderate pathological alterations in DLPFC/BA9 are implicated in PD and AD-related cognitive impairment [30, 31]. In this sense, analysis of tissue-specific neuropathological hallmarks obtained from high-throughput RNA-seq studies of post-mortem DLPFC/BA9 samples may provide insights related to early neurodegenerative mechanisms involved in AD and PD.

Given the relative inaccessibility of the brain during a person's lifetime, reliable, minimally invasive, and inexpensive biomarkers that accurately reflect the state and progression of brain diseases and the responses to potential treatments are regarded as the 'holy grail' in the clinical management of NDs [32]. This has led many to investigate peripheral biomarkers for detecting the earliest biological signals to identify patients at risk of developing NDs.

Brain-derived biomarkers are typically present at relatively low concentrations in the blood because the BBB prevents the free passage of molecules between the CNS and blood compartments. However, researchers have described BBB dysfunction in patients with AD and PD, which would result in the CNS-blood exchange of proteins and other molecules [33, 34].

In previous transcriptomic studies, the crucial role of the peripheral blood or WB in the diagnosis of AD and PD has been identified [35-38]. These reported a variety of blood-based dysregulated genes associated with innate, humoral, and adaptive immune processes involved in AD and PD

pathology [35, 36]. Similarly, serum phosphorylated form of tau (p-tau217) and α -syn (p-Ser-129) are blood biomarkers associated with the progression of the two most common NDs [37, 38].

Additionally, the study of PBMCs, which consist of multiple cell subsets, may provide insight into the pathogenesis of AD and PD. PBMCs, a potential source of myeloid and lymphoid cells, have critical roles in immune response and metabolism [39]. Myeloid cells are comprised of monocytes and their descendants, along with granulocytes such as neutrophils and basophils, while lymphoid cells primarily include T cells, B cells and natural killer (NK) cells. Additionally, PBMCs share much of the non-synaptic biochemical environment of neurons and contain the full complement of epigenetic enzymes and machinery found in neurons [39]. Many studies have shown that the investigation of PBMCs can recapitulate the conditions of the surrounding tissue, thus, providing a suitable source of biomarkers for predicting the progression of NDs [40, 41, 42].

In the following sections, I delve into the existing brain and blood research literature. This exploration aims to underscore the presence of potential biomarkers tied to alterations in genetic variants and gene expressions. Recognizing these biomarkers is fundamental for understanding how to diagnose or track the progression of both AD and PD. This foundational knowledge will set the stage for the in-depth exploration of AD and PD diseases, their sporadic and familial forms, and the involved pathways.

1.3 Alzheimer's Disease (AD)

AD, the most common form of ND worldwide, comprises 60% to 80% of all dementia cases and affects an estimated 24 million people globally [43].

Late-onset AD (LOAD; onset around 65 years of age) is the most common form of dementia in the elderly and presents with heterogeneous phenotypes due to interactions between genetic and environmental factors affecting various pathways and processes [44]. Memory impairment is

usually the most prominent feature in the early stages due to medial temporal cortex involvement. With disease progression, language and/or visuospatial deficits become evident as the pathology propagates to the material and parietal cortices. Participation of the frontal cortex in the late stages results in severe impairment in most cognitive domains, including judgment, abstract or logical reasoning, and planning [45].

The leading genetic risk factor for LOAD is the apolipoprotein E (APOE) genotype (haplotype). APOE, which encodes the brain's major cholesterol transporter, has three common alleles: $\epsilon 2$ (8.4% estimated allele frequency in the population), $\epsilon 3$ (77.9%), and $\epsilon 4$ [46]. ApoE $\epsilon 4$, as the major known genetic risk factor for LOAD, suggests that cholesterol metabolism and trafficking may be central to AD pathogenesis [47]. Potential synergistic interactions between APOE $\epsilon 4$ and aging-related metabolic changes is associated with an increased risk of developing AD via a multitude of mechanisms, including the enhanced formation of A β fibrils by accelerating the initial seeding or nucleation of amyloid- β (A β) polypeptide deposition, increased hyperphosphorylated microtubule-associated tau protein (p-tau), reduced glucose metabolism, vascular, and mitochondrial function, along with neuronal loss [46, 48]. Additionally, disease-related alterations in the environment in which neurons exist provoke glial cells, including astrocytes and microglia, accompanied by low to moderate levels of inflammatory mediators, which releases cytokines and free radicals, associated with increased risk of developing AD in aging brain [49]. However, APOE $\epsilon 4$ is not present in all LOAD cases, prompting a search for other genetic risk factors [50].

Some patients may present atypical phenotypes, including early-onset autosomal dominant AD (EOAD; onset 35-60 years of age) or non-amnesic presentations such as posterior cortical atrophy [51]. In general, there are 2 types of inheritance patterns observed in EOAD: mendelian (mEOAD; sometimes referred to as familial EOAD) and nonmendelian (nmEOAD) patterns. mEOAD forms

are typically fully penetrant with an autosomal dominant inheritance pattern caused by mutations in genes coding for amyloid precursor protein (APP) and presenilin genes (PSEN-1, PSEN-2); encoding the catalytic subunit of the γ -secretase complex that generates amyloid beta ($A\beta$) peptides via cleavage of APP in combination with β -secretase [51, 52]. Mutation of these genes results in increases in the ratio of $A\beta_{1-42}$ / $A\beta_{1-40}$ [44]. Conversely, nmEOAD is often sporadic or with inconsistent inheritance patterns. Although the aetiology and genetic basis of nmEOAD are mainly unknown, it is generally considered multifactorial, with polygenic effects [51].

In both EOAD and LOAD, $A\beta_{1-42}$ accumulates (soluble intracellular (monomer to oligomers) and insoluble extracellular (proto-fibril to fibrils)), involves neuropathological changes that initially affect neocortical regions, then the limbic system, diencephalon, basal forebrain, and lastly the cerebellum that leads to similar clinical phenotypes presentation in both forms [53, 54].

However, LOAD is immensely complex, and subsequently the current lack of effective treatment and the increasing aging of the population has transformed LOAD into a health and socioeconomic problem. Therefore, exploring molecular and cellular processes associated with disease progression in human AD central brain and peripheral blood provide valuable insights for better understanding of misfolded proteins pathological functions and their-mediated signalling cascades, useful for the development of practical biomarkers and more effective treatments for LOAD.

1.3.1 Propagation of $A\beta$ proteinopathy in LOAD

$A\beta$, the primary component of plaques in AD brains, is generated through sequential cleavages of APP by proteases. Understanding APP trafficking patterns within the endo-lysosomal system and the underlying mechanisms for $A\beta$ production are critical for elucidating the mechanisms of AD pathogenesis.

1.3.1.1 APP structure and endo-lysosomal trafficking

APP gene, a type-I integral transmembrane glycoprotein, is located on chromosome 21 in humans and belongs to a small gene family of APP-like proteins, including APLP1 and APLP2 [55]. APP has several isoforms generated by alternative splicing. Still, the three major isoforms are the 695 amino acid isoform (APP695), predominantly expressed in neurons, 751 and 770 amino acid isoforms (APP751, APP770) are expressed in most tissues and glial cells [55, 56].

APP is synthesised in the endoplasmic reticulum (ER), and its structure (Figure 1.4) includes a small intracellular C-terminal domain, an A β peptide region and a large extracellular N-terminal domain [55].

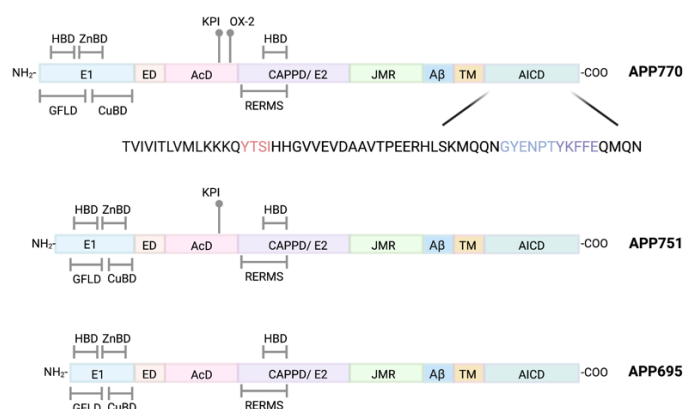


Figure 1.4 Domain architecture of the amyloid precursor protein (APP695).

Several domains, subdomains and functional segments have been described for APP architecture [57]. The N-terminal half includes the E1-domain, the cysteine-rich domain, the growth-factor-like domain (GFLD), the copper-binding domain (CuBD), a heparin-binding domain (HBD), a zinc-binding domain (ZnBD), the extension domain (ED), and the acidic domain (AcD) [55, 57]. Similarly, different domains are specified for the C-terminal half of the large APP-ectodomain, including the central APP domain (called CAPPD or E2-domain), a second heparin-binding domain, the RERMS-domain, one collagen-binding domain, the juxta membrane region (JMR) adjacent to a single transmembrane helix (TM) and the intracellular domain of APP (AICD). The longer splicing forms of APP (APP751 and APP770) also contain additional regions, the Kunitz-type-inhibitor (KPI) domain and the Ox-2 sequence in-between the AcD and E2 [55-57].

The endo-lysosomal trafficking pattern of APP is determined by its interactions with other proteins through its tyrosine-based motifs located in the C-terminal half: *YTSI*, located 4-amino acid residues away from the transmembrane domain, as well as *GYENPTY*, and *YKFFE* (Figure 1.4) situated within the AICD region [58]. From the ER compartment, the synthesised APP is transported to the Golgi/trans-Golgi network (TGN) apparatus, where it undergoes post-translational modifications such as N- and O-linked glycosylation, phosphorylation, and tyrosine sulphation during maturation [59]. Most APP localised in the TGN participates in multiple cellular events, including cell adhesion, neurite outgrowth, neuronal migration, protein transport, cell signalling and synaptogenesis [60]. Only a small proportion of APP is detected at the plasma membrane (PM), and over 50% is internalised within 10 minutes and sorted into early endosomes, where one fraction of APP is recycled back to the PM. Another fraction is targeted to the lysosome for degradation [59, 60]. The *YTSI* motif is involved in mediating endocytosis, while the *GYENPTY* motif is specifically involved in the re-internalisation of APP from the PM through its interaction with adaptor protein AP-2 via Disabled-2 (Dab2), a protein containing a phosphotyrosine-binding (PTB) domain [58]. The *YKFFE* motif is involved in the TGN-to-endosome transport of APP through binding with adaptor AP-4. Adaptor AP-3 also mediates APP trafficking from the Golgi to the lysosomes by binding to the *YTSI* motif [58].

Another point of interest stems from the fact that APP proteolytic processing is mediated with three proteases (α -, β - and γ -secretases), and understanding their trafficking pattern in relation to APP cleavage may therefore provide important clues for understanding AD pathology.

1.3.1.2 α -secretases endo-lysosomal trafficking

α -secretases are members of the disintegrin and metalloprotease family (ADAM10, ADAM17, and ADAM19), which comprises transmembrane and secreted metalloproteinases with functions

in cell adhesion and proteolytic processing of diverse cell surface receptors ectodomains and signalling molecules [61]. ADAM10 shows coordinated expression with APP in the human brain, which is less seen for ADAM17, and ADAM19 [61, 62]. ADAM10 is a multi-modular transmembrane protein expressed in neurons, vascular cells, leukocytes, and tumour cells [62]. It is synthesised in the ER as an inactive zymogen containing C-terminal domain, a transmembrane domain, a cysteine-rich domain, disintegrin and zinc-binding metalloprotease domain. There is also a pro-domain, which maintains the protease in a latent inactive form via a cysteine switch mechanism. A cysteine residue in the pro-domain holds a zinc ion in the catalytic site, preventing proteolytic activity [61, 62]. During transport to the PM, ADAM10 is N-glycosylated at four positions and undergoes maturation through pro-domain removal by pro-protein convertases. This is consistent with the notion that α -secretase proteolytic processing of APP occurs at the PM [62].

1.3.1.3 β -secretase endo-lysosomal trafficking

β -secretase or β -Site APP cleaving enzyme 1 (BACE1), a type-I transmembrane aspartyl protease with 501 amino acids, is synthesised in the ER. While trafficking through the Golgi apparatus, BACE1 undergoes maturation that involves cleavage of the pro-sequence, glycosylation and palmitoylation [63]. Mature BACE1 is sorted in the Golgi by a Golgi-localised, gamma adaptin ear-containing, ARF-binding 1 (GGA1) for the transport to the endosome, which is BACE1's primary location and site of activity [64]. A small proportion of BACE1 can traffic from the TGN to the PM, which concentrates as a dimer in lipid rafts. Then, it undergoes internalisation into specialised early endosomes under the control of GGA1 and ADP ribosylation factor 6 (ARF6) [63, 64]. BACE1 located in the endosome, can undergo ubiquitination, which is the recognition signal by GGA3 for its export to the lysosome and degradation. The acidic environment of the

TGN and endosome is optimal for BACE1 activity [64]. Thus, β -cleavage of APP primarily occurs in the TGN and endosomal/lysosome organelles.

1.3.1.4 γ -secretase endo-lysosomal trafficking

γ -secretase is an intra-membrane-aspartyl protease involved in cleavage multiple signalling proteins within their transmembrane domains (TMDs). It comprises a protein complex, including presenilins (either PSEN1 or PSEN2), nicastrin (NCT), anterior pharynx defective (APH)-1a or APH-1b, and presenilin enhancer-2 (PEN-2) [65]. PSEN1 is the core catalytic subunit of γ -secretase, in which two aspartate residues (Asp257 and Asp385) located in TMD6 and TMD7 are crucial for the catalytic activity of the enzyme. In the secretory pathway, PSEN1 undergoes endo-proteolysis, generating stable N-terminal and C-terminal fragments (PSEN1-NTF and PSEN1-CTF, respectively). The three other components (NCT, APH-1 and PEN-2) act as scaffold proteins to assemble and activate the γ -secretase protease complex [65].

Shortly after its synthesis, NCT is rapidly N-glycosylated to form the immature, partially glycosylated form (imNct). imNct binds to APH-1 in the ER to form a stable 1:1 complex. This stable heterodimer then binds the PSEN holo-protein in the ER to form the imNct·APH-1·holoPSEN heterotrimer [66]. The mNct·APH-1·holoPSEN trimer can now bind PEN-2. Assembly of all four components and the resulting endo-proteolysis allow the complex's exit from the ER to the Golgi via the ER-Golgi intermediate compartment (ERGIC). The majority (~95%) of mature γ -secretase complexes cycles between the ER and the Golgi, with a minority (~5%) further trafficked to the PM or other membranous components (e.g., endosomes, multivesicular bodies (MVBs), or lysosomes) [66]. Therefore, functional γ -secretase activity on APP usually occurs in the TGN or the early endosome domains.

1.3.1.5 APP proteolytic processing and A β production

Two different pathways can metabolise complex sequential proteolytic processing of APP, the non-amyloidogenic and amyloidogenic pathways, that lead to the production of several biologically active fragments that may have specific and even opposing functions [67]. In addition, two recent APP processing pathways have been identified, namely the δ -secretase (asparagine endopeptidase, AEP) and η -secretase (matrix metalloproteinase, MT5-MMP), capable of generating different fragments of extracellular and intracellular peptides [68].

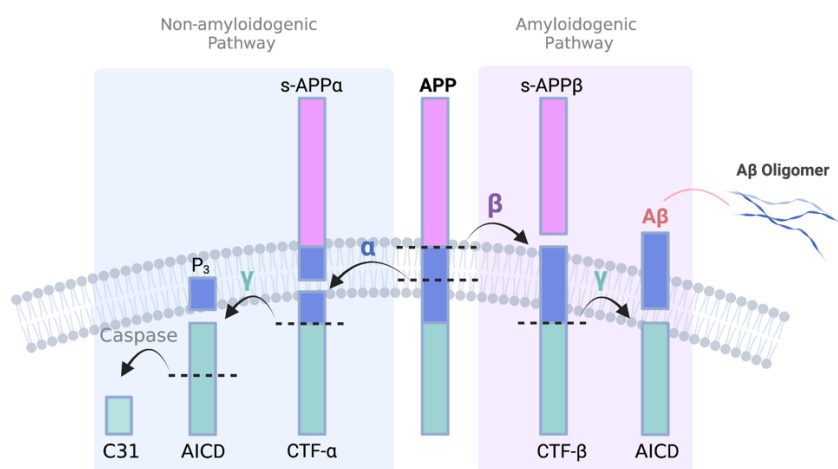


Figure 1.5 APP processing via the non-amyloidogenic and amyloidogenic pathways.

In the non-amyloidogenic pathway, APP is cleaved by α -secretase to form CTF- α , and an s-APP α . Cleavage of CTF- α via γ -secretase yields a short peptide fragment, P₃, and AICD. In the amyloidogenic pathway, APP is cleaved by β -secretase, which produces s-APP β and CTF- β fragments. CTF- β fragment is cleaved by γ -secretase to produce A β and AICD. A β further forms A β oligomers which subsequently results in the formation of fibrils and neurotoxic A β plaques extracellularly.

In the non-amyloidogenic pathway, α -secretase cleaves APP within the A β region, preventing the formation of A β and liberating a soluble ectodomain of APP (s-APP α) [69] (Figure 1.5). The remaining C-terminal fragment (CTF- α) or C83 is cleaved by γ -secretase, releasing the non-toxic peptide P₃ and AICD [69, 70] (Figure 1.5). In normal conditions, detecting AICD in cells and human

tissues is problematic because it undergoes rapid inactivation by the mainly cytosolic and endosomal insulin-degrading enzyme, a metalloproteinase, and by caspases to yield a peptide called C31 [71]. High levels of AICD appear to play an essential role in the onset of AD by tau phosphorylation/ aggregation and decreased cell proliferation/survival [72].

Concerning APP crowding, it has previously been shown that APP aggregation on PM is a prerequisite for endocytosis and amyloidogenic processing. More efficient crowding promotes internalisation and consequently diminishes the possibility of APP processing by α -secretases [73]. In the amyloidogenic processing pathway, internalised APP is cleaved by BACE1 and γ -secretase, mainly in the TGN and endosomes. β -secretase/ or BACE1 cleaves APP within the extracellular domain at position 671, shedding off a soluble ectodomain of APP (s-APP β) [69, 70] (Figure 1.5). The remaining C-terminal fragment (CTF- β)/ or C99 is subsequently processed by γ -secretase, releasing A β fragments of different lengths, including fragments of 43, 45, 46, 48, 49 and 51 amino acids that are further cleaved to the main final A β forms, mainly consist of 40 (A β ₁₋₄₀) or 42 (A β ₁₋₄₂) amino acids. As with the non-amyloidogenic pathway, the AICD fragment is believed to be secreted into the extracellular space [72].

1.3.2 A β function under homeostatic and pathological conditions

A β monomers produced under homeostatic conditions can trigger or sustain intracellular signalling essential for neurotransmission (including the regulation of the excitation/inhibition balance and synaptic vesicle trafficking) and synaptic plasticity by mediating the cyclic adenosine monophosphate response element-binding protein (CREB)-mediated transcription of the brain-derived neurotrophic factor (BDNF) axis pathway [74]. A β monomers are also proposed to have multiple roles in neurogenesis, metal ion sequestration, antioxidant activity, and calcium homeostasis. In healthy individuals, the activation of A β monomers is observed in microbial and

virus infections. Microbial infection studies revealed a protective role for A β ₁₋₄₂ in innate immunity as an antimicrobial peptide (AMP), characterised by reduced microbial adhesion to host cells, induced agglutination and finally, entrapment and perforation of microbes by A β fibrils [75-77]. A β ₁₋₄₀ antimicrobial activity is not as potent as A β ₁₋₄₂ and is only known for its slight antifungal activity. In virus infections, A β ₁₋₄₂ peptide establishes interactive bonds with the viral coat proteins, which induces their aggregation and prevents their fusion and/or uptake by target cells, leading to an innate recognition of virus particles by neutrophils and, finally, their clearance [78]. Influenza A and herpes simplex virus type 1 (HSV1) are among the viruses strongly inhibited by A β protective activity [79].

However, under pathological conditions, the production of A β increases, further promoting the accumulation and deposition of A β into soluble oligomers, which then form insoluble fibrils that are finally deposited as amyloidogenic plaques [80].

A β oligomers can directly act on the cell membrane to form channels and destroy membrane integrity and permeability, rendering neurons vulnerable to excitotoxicity and apoptosis [81]. Simultaneously, the interaction of A β oligomers with the receptors on the surface of neurons, such as cellular prion protein (PrP^c), glutamate receptors (N-methyl-d-aspartate receptor (NMDAR), and α -Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPA)), β 2-adrenergic receptor (β ₂-AR), ephrin type-B receptor 2 (EphB2), ephrin type-A receptor 4 (EphA4), p75 neurotrophin receptor (p75NTR), and α 7-nicotinic acetylcholine receptor (α 7nAChR) can alter downstream signalling pathways leading to A β mediated synaptotoxicity and ultimately neuronal cell death [82]. A β oligomers binding to the extracellular N-terminus of β ₂-AR interfere with AMPAR activity by inducing G-protein/cAMP/protein kinase A (PKA) signalling, which dysregulate intracellular Ca²⁺ homeostasis leading to mitochondrial Ca²⁺ overload, activation of

mitochondrial permeability transition pores (MPTP), and an increase in Ca^{2+} -dependent proteases (calpains) activity that directly induce reactive oxygen species (ROS) generation, and ion-motive ATPases inhibition leading to synaptic dysfunction, and neuron loss [83].

In addition, A β oligomers can cause a chronic response of the innate immune system by activating microglia through binding to immunological receptors such as toll-like receptors 2 (TLR2), TLR4, TLR6 and their co-receptors, including CD14, CD36, and CD47 [84]. Furthermore, A β oligomers can lead to impaired mitochondria function and dynamic possibly through interactions with mitochondrial proteins, such as adenine nucleotide translocase (ANT), components of the translocase of the outer membrane (TOM) or translocase of the inner membrane (TIM), dynamin-like protein-1 (DLP1), cyclophilin D, and uncoupler protein (UCP) [85]. Moreover, the accumulation of A β oligomers led to intracellular tau proteins phosphorylation and aggregation, another characteristic feature of AD pathogenesis, discussed in detail below.

1.3.2.1 A β deposits and induced-tau pathology

Tau protein is a neuronal protein with 6 isoforms generated through alternative splicing of exons 2, 3 and 10 of the microtubule-associate protein tau (MAPT) gene. In the adult human brain, tau isoforms range from 352 to 441 amino acids in length (a longer isoform is predominant in peripheral tissues), and their structural sequence is divided into N-terminal region, proline-rich region (PRR), microtubule-binding domain (MTBD), and C-terminus region [86] (Figure 1.6 A).

In axons, the tau protein's "free" flanking N-terminal and C-terminal regions interact with two motor proteins (kinesin and dynein) involved in facilitating the axonal transport of cellular cargoes on microtubule tracks [87]. In synapses, the tau protein binds actin via its PRR domain and, at the same time, also binds microtubules through its MTBD and serves as a crosslinker between microtubules and actin filaments and thus helps to organise the cytoskeleton network [88].

Additionally, the PRR region has several threonine and serine residues (Ser202, Thr205, and Ser208), the phosphorylation of which by specific kinases has known physiological functions, including maintenance of microtubule assembly and stability [86, 89]. However, phosphorylation of monomeric tau beyond specific thresholds can form neurotoxic oligomeric aggregates, paired helical filaments (PHFs), and ultimately neurofibrillary tangles (NFTs) that have pathophysiological consequences in AD [89]. Simultaneously, disruption of the hyperphosphorylated tau degradation mechanism, including ubiquitin-proteasome system (UPS), chaperon-mediated autophagy (CMA), and endosomal micro-autophagy (eMI), can trigger intracellular tau accumulation and downstream tau-dependent neuronal dysfunction [90].

Hyperphosphorylated tau can propagate from one neuronal cell to another via direct translocation (Figure 1.6 (B1)), secretion (Figure 1.6 (B2)), receptor-mediated endocytosis or through extracellular vesicles (EVs), including exosomes (Figure 1.6 (B3)) and ectosomes (Figure 1.6 (B4)), bulk endocytosis (Figure 1.6 (B5)) and/or by tunnelling nanotube (TNT) (Figure 1.6 (B6)), associated with enhanced aggregation of intracellular tau and increased risk of AD [91].

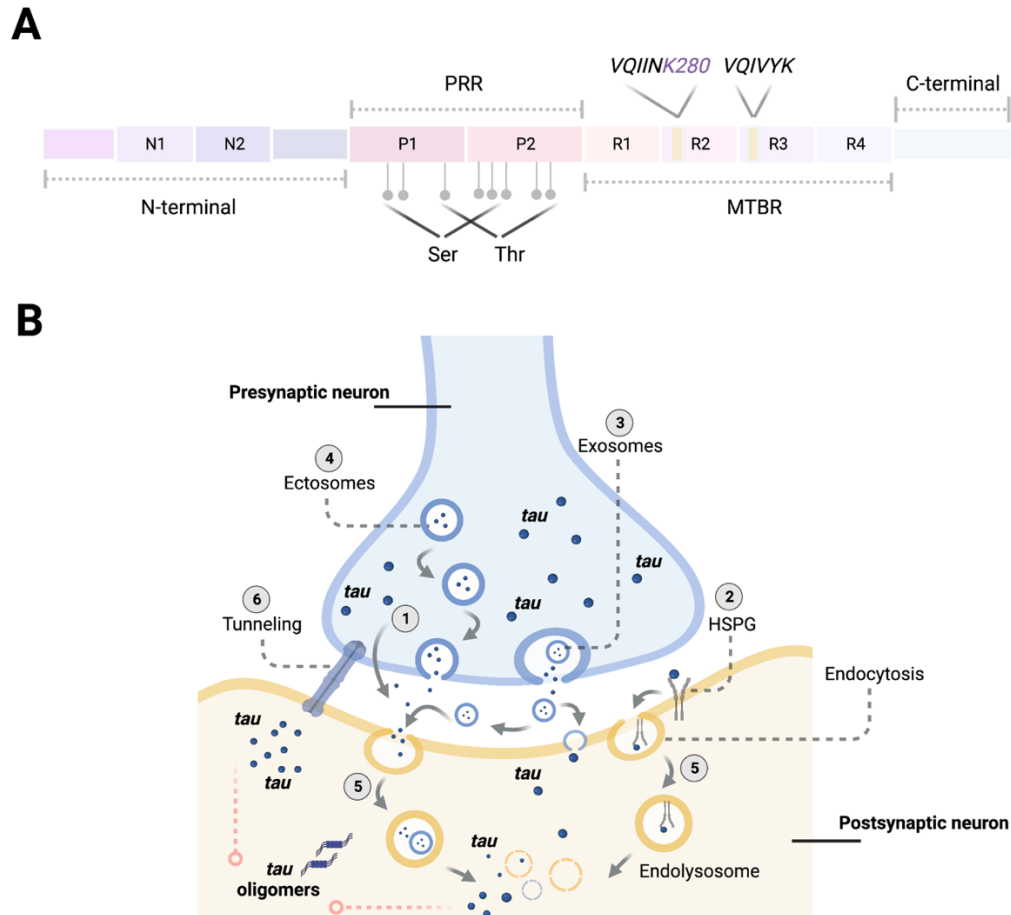


Figure 1.6 Structure of the full-length tau protein and molecular mechanisms for cell-to-cell transmission of pathogenic tau.

A) Schematic showing tau (441 amino acids), the longest isoform expressed in the human brain. Tau protein contains major structural domains, including N-terminal domain with N1 and N2 inserts, a proline-rich region (PRR) with threonine and serine phosphorylation sites, four major microtubule-binding repeats (MTBR; (R1-R4)), and C-terminal domain. Hexapeptide *VQIINK280* motif in the second repeat and *VQIVYK* in the third repeat within the tau MTBR are crucial for forming β -sheet structures and consequent tau aggregation. **B)** Schematic representation of molecular mechanisms for the transmission of pathological tau between donor neuronal cells (shown in blue) and acceptor neuronal cells (shown in yellow). **1)** direct translocation of pathological tau, **2)** pathological tau binds heparan sulfate proteoglycans (HSPGs) on the cell surface and then taken up via macropinocytosis, a form of endocytosis, **3)** packaging of pathological tau in exosomes by inward budding of late endosome membrane leading to the formation of intraluminal vesicles in MVBs that can be secreted by fusion of MVB membrane with presynaptic neuron, **4)**

packaging of pathological tau in ectosomes. Ectosomes are larger than exosomes and differ in their molecular composition. After their release from cells, both ectosomes and exosomes function similarly and can be fused to or endocytosed by target cells (postsynaptic neurons), **5**) endocytosis is a process by which cells absorb and engulf large molecules such as proteins that generally cannot cross the hydrophobic plasma membrane. It has been proposed for uptake of multiple exogenous aggregates, including tau, **6**) transfer of pathological tau via tunnelling nanotubes that directly connect the cytosols of two neighbouring cells.

Current disease models suggest that A β , either as plaques or as nonfibrillar, and oligomeric forms, synergistically accelerates tau pathology (tau misfolding and aggregation) that spreads throughout the cortex and arguably leads to tau-induced toxicity, resulting in neural system failure neurodegeneration, and cognitive decline [92]. In this context, researchers have proposed a link between extracellular A β deposition and induction of Ca²⁺-influx into neurons, leading to downstream Ca²⁺-elevations that can aberrantly stimulate signalling cascades, including the activation of protein kinases (FYN, cyclin-dependent kinase-5 (CDK5), Glycogen synthase kinase-3 beta (GSK3- β), and p38 mitogen-activated protein kinase (MAPK)), and calcineurin mediated pathways which affect tau hyperphosphorylation [93]. A β has also been shown to directly induce the calpain-mediated cleavage of tau, leading to the generation of a toxic 17-kDa fragment and inducing abnormal tau phosphorylation at disease-relevant sites [94]. The abnormally phosphorylated tau then dissociates from the microtubules. It accumulates at the dendritic compartments, where it may affect the synaptic trafficking and/or anchoring of glutamate receptors (glutamate ionotropic receptor AMPA type subunit 2 (GRIA2/ or GLUR2)), which in turn affects postsynaptic function leading to mitochondrial dysfunction, synaptic deficits, and memory impairment [95]. Additionally, hyperphosphorylated tau can interact with FYN leading to induced NMDAR sensitise, which renders neurons more vulnerable to the toxicity of A β oligomers in the postsynaptic compartments [82].

1.3.3 Clearance or removal of A β deposits

To control diverse A β functions and to prevent neurotoxic accumulation of A β in the brain, A β products must be cleared by several clearance pathways (Figure 1.7). Recent evidence in patients with APOE ϵ 4 gene status has suggested that faulty A β clearance mechanisms along with overproduction contribute to significantly higher levels of monomeric A β_{1-40} and A β_{1-42} that can aggregate into different assemblies giving origin to oligomers and amyloid fibrils that are insoluble and can further aggregate into cerebral A β plaques [96, 97]. Pathways crucial for the clearance of pathogenic A β are represented below.

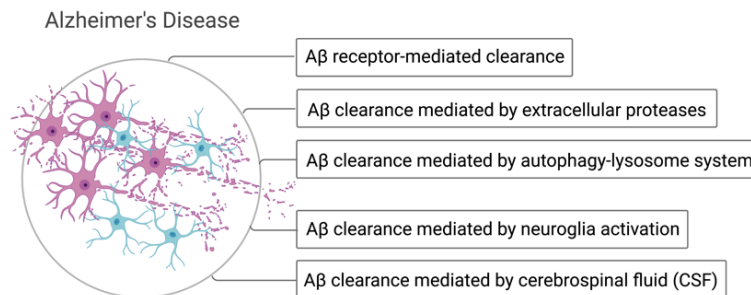


Figure 1.7 Pathways associated with A β clearance in the brain.

1.3.3.1 A β receptor-mediated clearance

In healthy brains, ApoE ϵ 4 regulates A β accumulation by stimulation of APP endocytosis and processing via ApoE receptor 2 (ApoER2) or lipoprotein receptor-related protein 8 (LRP8) [98]. ApoER2 interaction with its ligand, the reelin protein, drives sequential proteolytic processing, resulting in the cleavage of the receptor by α -secretase, which generates a membrane-tethered C-terminal fragment (ApoER2-CTF), followed by the cleavage by γ -secretase [99]. The action of γ -secretase generates an intracellular domain fragment (ApoER2-ICD). ApoER2-ICD can regulate the expression of low-density lipoprotein receptors (LDLR) [99]. Several studies have indicated that A β clearance are modified by the LDLR family consisting of several cell surface receptors,

including the LDLR-related protein 1 (LRP1), LRP5, LRP6, ApoER2/LRP8, LRP2LDLR-related protein 1B (LRP1B), very low-density lipoprotein receptor (VLDLR), and sortilin-related receptor 1 (SORL1) [47, 99]. LRP1 is primarily expressed in the CNS and located at the abluminal surface of the brain-facing side of polarised endothelial cells [100]. LRP1 is composed of two subunits that include an extracellular α -chain (515 kDa) and a β -chain (85 kDa) containing a short extracellular extension, a single transmembrane domain and a 100 amino acid intracellular domain [97, 100]. The intracellular subunit also contains two *NPxY* motifs (*YXXL* motif and di-leucine motifs) important for adaptor protein binding, signalling and endocytosis [100]. LRP1 promotes A β clearance by binding to A β through the adaptor protein amyloid beta (A4) precursor protein-binding, family B, member 1 (APBB1/Fe65) and promotes rapid A β endocytosis to early endosomes [59]. LRP1 then associates with phosphatidylinositol binding clathrin assembly (PICALM) protein that guides trafficking of A β -LRP1 endocytic vesicles to Rab5-positive early endosomes and then to Rab11-positive sorting endosomes, handing A β over to P-gp transporter (also known as ATP binding cassette subfamily B member 1 (ABCB1)), which then pumps it out of the luminal side of endothelial cells directly into the blood [47, 100, 101]. Thus, the downregulation of cell surface LRP1 significantly decreases A β clearance from the brain of AD patients and aging adults.

1.3.3.2 A β clearance mediated by proteases

In addition, a growing body of evidence suggests that A β released into the extracellular space can be degraded by proteases secreted from neurons and glial cells [72]. Neprilysin (NEP), insulin-degrading enzyme (IDE), angiotensin-converting enzyme (ACE), endothelin converting enzymes 1 and 2 (ECE1 and ECE2), and matrix metalloproteinase-2 (MMP-2), MMP-3, MMP-9, are the most efficient A β proteases that are reported to have an affinity for specific domains within the

soluble A β ₁₋₄₀ and A β ₁₋₄₂ amino acid sequence to convert these peptides into non-toxic fragments [102]. Genetic and physiological studies have shown that changes in the expression and activity of these proteases resulted in elevated A β accumulation and cognitive impairment.

1.3.3.3 A β clearance mediated by autophagy-lysosome system

Autophagy is a major intracytoplasmic protein degradation pathway in which cytoplasmic contents are trafficked along microtubules by double-membraned vesicles called autophagosomes (AP) to enable the AP fusion with the lysosome, which is mainly located around the nucleus to form autolysosomes (ALs) [103]. Monolayer membranes ALs facilitate the degradation of cellular substrates, which may include A β and APP. However, A β may also be generated in AP, which contains both APP, β -secretase and γ -secretase complexes. Typically, APs can be quickly transported to the cell body and degraded, so APs are rarely seen in normal nervous tissue [104].

In AD, autophagy is markedly impaired, evidenced by the considerable accumulation of unprocessed A β -rich APs in dystrophic neurites [105]. Reduced expression of key autophagic proteins (beclin-1 and autophagy-related (ATG) protein 5 and 7 (ATG5 and ATG7)) involved in membrane elongation and engulfment of a variety of substrates (e.g., protein aggregates, infectious agents, and damaged mitochondria) along with the up-regulated mammalian target of rapamycin complex 1 (mTORC1) signalling pathway likely contribute to autophagic dysfunction, A β accumulation, and neuronal cell death [102]. mTORC1 is required for autophagy inhibition by binding to the Unc-51-like kinase 1 (ULK1) complex and phosphorylating both ATG13 and ULK1, thereby suppressing ULK1 kinase activity and preventing the interaction between ULK1 and AMP-dependent protein kinase (AMPK) which promotes autophagy [103].

In addition, mutations in specific lysosomal ion channels, including transient receptor potential channel mucolipin (TRPML) protein family, two-pore channels (TPCs) and the chloride channels

(CLCs)), and proton pumping vacuolar H⁺-ATPase (or V-ATPase) have been associated with changes in lysosome membrane permeability to several ions (H⁺, Ca²⁺, Na⁺, K⁺, and Cl⁻) and cations (Ca²⁺, Fe²⁺, Na⁺ and Zn²⁺) contributing to lysosomal enzymatic defects, pathogenic A β proteins aggregation and uncontrolled neuronal cell death [106].

1.3.3.4 A β clearance mediated by neuroglia activation

The glia are the non-neuronal cells of the central nervous system. They are classified based on their morphology, functions, and location in the nervous system. The crosstalk between different glial cells and neurons provides homeostatic, structural, and metabolic support of brain cells, which are crucial in brain development and function. Astrocytes, microglia, and oligodendrocytes (ODs) are the principal glial cell types of the CNS [107]. Microglia cells are thought to synchronise with the development, maturation, and senescence of the CNS throughout life via the adoption of different regulatory networks and have essential roles in developmental synaptic pruning, maintenance of synaptic plasticity, neuronal apoptosis, and immune surveillance [108]. In comparison, astrocytes are characterised by their critical roles in BBB structure, CNS homeostasis, synapse formation, strength, and myelination [109].

Microglia express many immunological receptors, such as TLR2, TLR4 and TLR6, and their co-receptors, including B-class scavenger receptors CD36, CD14, CD47 and α 6 β 1 integrin [110]. Specific interaction of TLR2, TLR4 and CD14 with fibrillar A β lead to microglial activation of p38-MAPK, which is required for fibrillar A β -induced ROS production and phagocytosis of fibrillar A β [110]. However, if the A β concentration rises above a critical threshold, oligomeric or fibrillar forms of A β are recognised as danger-associated molecular patterns (DAMPs) by CD36, which triggers the formation of a TLR4-TLR6 heterodimer, leading to activation of nuclear factor- κ B (NF- κ B). NF- κ B activation results in the upregulation of downstream target genes encoding

components of the NOD, LRR- and pyrin domain-containing 3 (NLRP3) inflammasome and proinflammatory cytokines, including pro-interleukin-1 β (pro-IL-1 β) [111]. NLRP3 inflammasome is a multimeric protein complex consisting of the danger sensor protein NLRP3, a signal adaptor protein known as apoptosis-associated speck-like protein containing a CARD domain (ASC), and the cysteine protease, caspase-1 [90]. Misfolded protein aggregation, including A β aggregates and environmental/ cellular stress, triggers NLRP3 activation and oligomerisation, leading to ASC recruitment and polymerisation and caspase-1 recruitment, dimerisation, and activation [112, 113]. Within the assembled inflammasome, active caspase-1 cleaves pro-IL-1 β into its bioactive form (IL-1 β) that can be released into the extracellular space, propagating the chronic neuroinflammatory signal to neighbouring cells [113]. IL-1 β secretion by microglia can also exacerbate tau phosphorylation and tangle formation through aberrant activation of p38-MAPK and GSK3- β , which affect synaptic plasticity, inhibiting long-term potentiation (LTP), and subsequently learn and memory [114].

1.3.3.5 A β clearance mediated by cerebrospinal fluid (CSF)

CSF is a clear fluid that fills a large space within and around the CNS and is produced in a highly vascularised portion of the ependymal lining of the lateral ventricles called the choroid plexus (ChP). The ChP is comprised of vascular stroma covered by a monolayer of epithelial cells containing a variety of transporters, membrane proteins, and immune cells that regulate this epithelial blood-CSF barrier (BCSFB) [115]. The ChP passively filters plasma through its capillary endothelial cells and stroma, after which the BCSFB epithelium actively filters and secretes ions into a modified plasma that is pushed along an osmotic gradient to become the intraventricular CSF [115]. CSF produced within the ventricular ChP flows into the subarachnoid space surrounding the brain and spinal cord and can either flow into the glymphatic system or be

absorbed via subarachnoid villi into the cerebral venous sinuses [116]. The glymphatic system is a brain-wide fluid transport pathway that works as a “cerebral lymphatic system” responsible for neuronal extracellular waste protein clearance via a paravascular pathway, mainly CSF-interstitial fluid (ISF) exchange [117]. The water channel aquaporin 4 (AQP4) in the vascular endfeet of astrocytes is required for rapid, high-volume CSF transport from the perivascular compartment into ISF space and waste drainage through the glymphatic system [117, 118]. In AD brain, changes in the perivascular AQP4 polarisation contribute to the impairment of glymphatic function, significantly impairing the paravascular CSF-ISF exchange associated with decreased clearance of A β aggregates, synaptic plasticity alteration and progressive cognitive decline [119, 120].

Based on the studies discussed here, the imbalance between production and clearance of A β_{1-40} A β_{1-42} in the early stages of AD is assumed to be a causative and initiating factor of this disease. Subsequently, AD progression is mediated by aggregation of A β monomers into soluble A β oligomers that trigger a series of toxic reactions in neurons, such as receptor disability, mitochondrial damage, Ca²⁺ homeostasis dysregulation, neuroinflammation, and abnormal tau phosphorylation accompanied by progressive loss of synapses, which results in gradual cognitive impairment and, finally, dementia.

1.4 Parkinson’s Disease (PD)

PD is the second most common ND after AD, affecting 6.1 million people worldwide. The prevalence of PD is estimated to be 0.3% in the general population, ~1% in people older than 60, and ~3% in people aged 80 years or older [2]. Pathological characteristics of PD are associated with the loss of dopaminergic (DA) neurons in the substantia nigra pars compacta (SNpc) of the midbrain, decreased release of dopamine, and the presence of insoluble proteinaceous cytoplasmic inclusions termed Lewy bodies (LB) and Lewy neurites (LN) in the remaining neurons [121]. Most

cases of PD manifest in a late-onset sporadic manner [122]. However, autosomal early-onset PD occurs only in a small number of families and is attributable to mutations in specific genes, including; duplication, triplication or point mutations in the α -synuclein gene (SNCA) and PD-like mutations in parkin RBR E3 ubiquitin protein ligase (PRKN), Parkinson's disease protein 7 (PARK7, also known as DJ-1), PTEN-induced kinase 1 (PINK1), leucine-rich repeat kinase 2 (LRRK2), and glucocerebrosidase (GBA) [122, 123].

The symptoms in both type of PDs is characterised by motor symptoms such as bradykinesia/akinesia, resting tremor, rigidity, and postural abnormalities, and non-motor symptoms such as dementia, hyposmia, depression, and emotional changes [53]. Studies in animal models and human post-mortem tissues have identified distinct cellular and molecular changes associated with PD pathology, and this is elaborated on below.

1.4.1 Propagation of α -synucleinopathy in PD

The pathogenic mechanisms involved in familial and sporadic forms of PD are likely to be shared. Both PD types are characterised by progressive accumulation of pathological α -syn inclusions in the brain. Therefore, α -syn biology and its abnormal structures are thought to be central to the initiation and development of PD.

1.4.1.1 α -syn structure

α -syn is a 15 kDa intracellular protein encoded by the SNCA gene localised on the long arm of chromosome 4 and enriched in many compartments, such as cytosol, mitochondria, pre-synaptic nerve terminals, and nucleus [124]. This protein comprises 140 amino acids distributed over three domains (Figure 1.8). The first region, the N-terminal, is composed of 4 regions of 11 “imperfect” repeats with a lysine-rich consensus sequence (*KTKGEV*) [124]. This sequence is relevant for α -helix conformation in the α -syn protein, facilitating its binding to negatively charged lipid

membranes and vesicles [124, 125]. Then follows the central region, referred to as the “non-amyloidogenic component” (NAC), known for its hydrophobic properties that become prone to aggregation under specific conditions, being responsive to conformational changes in α -syn, from random coiled-coils to β -sheet structures in an oligomeric state [124, 125]. The last region is the C-terminal region, which is characterised by acidic residues implicated in regulating α -syn nuclear localisation and fibrillation, as well as interactions with metals, small molecules, and other proteins [125].

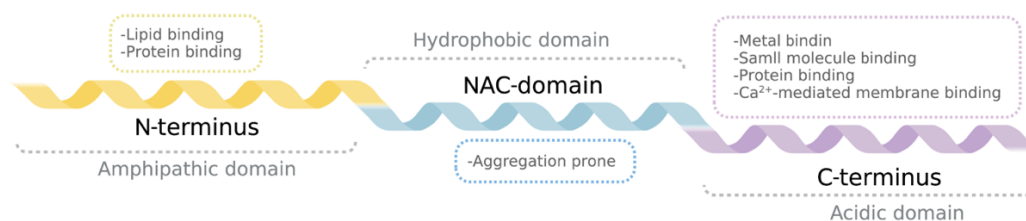


Figure 1.8 The structure of monomeric α -synuclein (α -syn).

Three regions define monomeric α -syn: N terminus (yellow), where the familial mutations leading to early onset Parkinson’s disease reside. NAC region (blue) and C terminus (purple), where many post-translational modifications occur.

1.4.1.2 α -syn structural plasticity and interactions

To understand α -syn functions in health and disease, different conformational states of α -syn have been studied. The structural plasticity of α -syn depends on cellular homeostasis and its interactions with other cellular components [124]. Cytosolic α -syn is natively unbound, unfolded, and monomeric. When bound to phospholipid membranes, α -syn takes on an α -helical conformation, allowing a more favourable binding of hydrophobic surfaces to the membrane's hydrophobic interior [126]. α -syn direct interaction with negatively charged phospholipid membrane/vesicles is crucial for regulating membrane stability, neuronal plasticity, synaptic rearrangement, controlling secretory and vesicular trafficking, neurotransmission, and influencing mitochondrial

fusion and fission, and lipid and calcium homeostasis [127, 128]. α -syn plays an essential role as a defence molecule in the innate immune system. α -syn can perform antimicrobial and antifungal activity against various bacterial and fungal strains through interactions with negatively charged lipopolysaccharide (LPS), peptidoglycan (PG), cardiolipin and phosphatidylglycerol of bacterial cytoplasmic membranes and sterol of fungal cell wall components [129].

However, in aged neurons with impaired degradation machinery, monomers of α -syn have been shown to take on a 'Greek-key'-like β -sheet topology and follow the nucleation-conversion-polymerisation aggregation model [130]. The kinetics of this pathway initiates with the primary nucleation phase, in which the native monomeric protein misfolds and combines to form 'pre-nucleation clusters' or 'pre-nuclei', which then convert into oligomer seeds (nuclei) that subsequently elongate via monomer addition, to give rise to fibrils. Upon maturation, the fibrils, which are otherwise less toxic, fragment into oligomers that act as secondary nuclei or 'seeds' for de novo fibrillation on existing fibril surfaces, giving rise to a feedback loop of fibril amplification to form inclusions [130].

The toxic aggregation of α -syn depends on specific changes in cellular microenvironment conditions (pH, temperature, ionic strength), macromolecular crowding, and the presence of other ligands/partners (amyloidogenic proteins, metal ions, intermediary toxic species, membranes, and specific lipid molecules) [131]. In addition, post-translation modifications (PTMs) of pathogenic α -syn can affect its size, charge, structure, and/or conformation, leading to α -syn aggregation and toxicity [132]. Interestingly, frequent cleavage of soluble α -syn at the C-terminal by Calpain1 promotes further aggregation of α -syn that potentially promotes the propagation of α -syn neuropathology [133]. Accordingly, A53T and A30P mutations in α -syn have consistently been

shown to increase the rate of α -syn self-aggregation by long-range interactions with the C-terminal residues, which contribute to autosomal dominant early-onset PD [134].

1.4.1.3 Transmission of toxic forms of α -syn in the brain

Several studies on cultured cells indicate that monomeric and oligomeric forms of pathogenic α -syn can be actively secreted in the extracellular space [135] by non-classical, ER/Golgi-independent protein export pathway [136]. Additionally, a fraction of α -syn was secreted in association with exosomes and other EVs through a calcium-dependent mechanism [135, 136]. Whichever mechanisms dictate structurally abnormal or damaged α -syn release, to achieve pathological intracellular actions, including seeding and prion-like spreading, α -syn must re-enter the cell from the extracellular space [133]. Endocytosis is the main pathway of internalising α -syn fibrils. A study demonstrated that α -syn fibrils bind HSPGs on the cell surface and are taken up via macropinocytosis, a form of endocytosis [137]. Another study showed that α -syn fibrils are recognised by lymphocyte activation gene 3 (LAG3) receptors on the cell surface, then internalised via receptor-mediated endocytosis [133, 137]. Internalised α -syn fibrils are then delivered to the endo-lysosomal pathway, where sorting occurs. These trafficking steps are highly selective and involve a series of membrane fusion/fission events mediated by the Rab family of small GTPases [138]. Most of the α -syn fibrils taken up by neurons are retained in lysosomes to be proteolytically processed. However, a small amount of misfolded α -syn ends up in the cytoplasm due to age-dependent lysosomal integrity impairment that could lead to pathogenic alpha-synuclein accumulation [139]. Aside from release and uptake mechanisms, TNT-mediated cell-to-cell transfer of pathological α -syn aggregates has also been reported. Through this mechanism, α -syn fibrils were incorporated in lysosomes of donor cells and then transferred through TNTs to

acceptor cells, subsequently forming inclusions by recruiting endogenous α -syn in acceptor cells [138].

Based on α -syn aggregation and transmission hypothesis, Braak and colleagues have proposed Braak staging for the progression of LB pathology in the brain of PD patients. PD pathology begins initially in the dorsal motor nucleus of the vagus nerve (dmX) and olfactory bulb (stage 1). The pathology moves from the dmX through the nuclei of the pontine tegmentum (stage 2) until it reaches the midbrain and, in particular, the substantia nigra (stage 3) [140]. From there, it passes through the temporal neocortex and allocortex (stage 4) to sensory association areas, the prefrontal cortex, and finally to the premotor cortex and primary sensory and motor cortices (stages 5, 6) [140, 141].

1.4.2 Misfolding and aggregation of α -syn: implications in PD pathogenesis

Abnormal misfolding and aggregation of α -syn is directly associated with PD pathogenesis through interactions with disease-causing factors that affect crucial cellular functions in the brain (Figure 1.9), which is discussed in detail below.

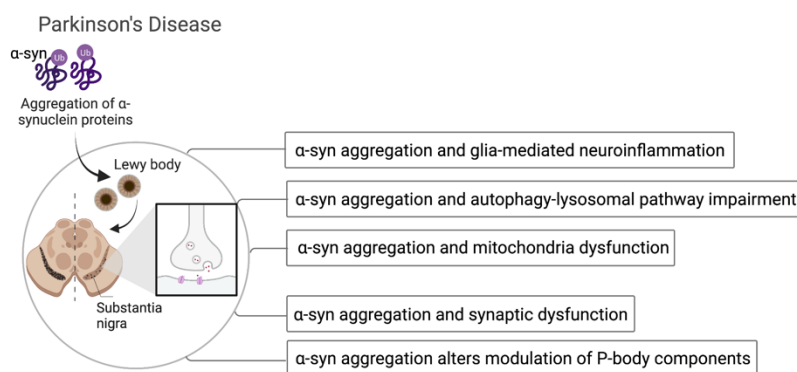


Figure 1.9 α -syn aggregation affecting critical cellular functions in the brain.

1.4.2.1 α -syn aggregation and autophagy-lysosomal pathway impairment

Preferentially, monomeric α -syn is processed for degradation by the autophagy-lysosomal pathway (ALP). CMA is an ALP that allows the degradation of α -syn by lysosome [142]. In this pathway, α -syn pentapeptide conserved motif biochemically related to *KFERQ* is selectively recognised by the cytosolic chaperone heat shock cognate protein of 70 kDa (hsc70). After binding to the chaperones, the substrate is targeted to the lysosomal membrane, interacting with the lysosome-associated membrane protein type 2A (Lamp2A). Lamp2A, in association with other proteins, forms a multi-protein complex and regulates the substrate translocation into the lumen of the lysosome, where the substrate will be degraded [142]. Additionally, oligomeric α -syn and various misfolded proteins are cleared via macroautophagy, following the fusion of autophagosomes with the lysosome.

When α -syn reaches toxic levels in disease conditions, it can interact with the transcription factor EB (TFEB), a master regulator of cellular clearance [143]. Usually, TFEB can enhance the autophagy-lysosome pathway, including lysosomal biogenesis, expression of autophagy genes and lysosomal proteostasis through its translocation to the nucleus and binding to the promoter region of the target genes [144]. Enhanced interaction of α -syn fibrils with TFEB would lead to TFEB sequester in the cytoplasmic inclusions and its activation repression, which results in lysosomal permeabilisation (LMP) due to inflammatory pathways activated by aggregated α -syn [144, 145]. This led to an enhanced secretion of toxic α -syn aggregates by neuronal cells in vitro and in vivo, causing inflammation and neurotoxicity.

1.4.2.2 α -syn aggregation and glia-mediated neuroinflammation

Monomeric or oligomeric α -syn released by neurons usually is taken up by microglial cells via clathrin-mediated endocytosis and is then driven to the lysosome for degradation. In pathological

conditions, α -syn aggregates may indirectly cause an increase in oxidative stress when it binds to microglial cell surface receptors TLR2, TLR4, and CD11b, resulting in the activation of inflammatory pathways, including NF- κ B and MAPK [146]. Induction of NF- κ B and other classical inflammatory pathways within microglia are also responsible for astrocyte activation [146]. In times of toxic insult (α -syn aggregates), increased release of proinflammatory cytokines such as tumour necrosis factor- α (TNF- α) and IL-1 β from microglia can induce the morphological and biochemical transformation of astrocytes, known as reactive astrogliosis [147]. Reactive astrogliosis is characterised by changes such as upregulation of glial fibrillary acidic protein (GFAP) and vimentin (Vim), increased expression of astrocytic glutamate transporters and potassium channels, decreased production of antioxidant molecules such as glutathione (GSH), and toxic release of ROS/ reactive nitrogen species (RNS) that lead to astrocytic oxidative/nitrosative stress, neuroinflammation, neuronal apoptosis, and PD pathogenesis [145, 147].

1.4.2.3 α -syn aggregation and mitochondria dysfunction

Mitochondria are present in the synaptic compartment and crucial for maintaining synaptic activity and neuronal homeostasis by ATP production, Ca²⁺ storage, and lipid metabolism [148]. Direct interaction of α -syn aggregates has been shown with voltage-dependent anion channel 1 (VDAC1) and inositol triphosphate receptors (IP3Rs), the protein components involved in mitochondria-associated membranes (MAMs) domains of ER-mitochondria contact, relevant for the exchange of ions and lipids [148]. The α -syn-MAMs interaction has been suggested to enhance mitochondrial Ca²⁺ uptake from the ER, which could, in turn, affect ATP production and oxidative stress [149]. Additionally, binding of α -syn aggregates to translocase of the outer membrane 20 (TOM20), a regulatory subunit of the TOM complex, has been shown to prevent TOM20

interaction with the co-receptor, TOM22, and result in impaired mitochondrial import of the NADH dehydrogenase [ubiquinone] iron-sulphur protein 3 (NDUFS3) subunit of complex I [150]. Linked to lower basal respiration, increased ROS and cytochrome C release leading to activation of caspase-3, caspase-9, and other pro-apoptotic factors ultimately leads to neuronal cell death [150].

1.4.2.4 α -syn aggregation and synaptic dysfunction

α -syn is important for the normal function and integrity of synapses. Pathologically, studies in mice have revealed an interaction of large α -syn oligomers with the N-terminus of multiple vesicle-associated membrane protein 2 (VAMP2/synaptobrevin-2) component of the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex [151]. This incapacitates VAMP2, a vesicle-associated (v)-SNARE ability to interact with syntaxin 1A (STX1A), associated with the PM ((t)-SNARE), inhibiting SNARE complex formation at the synaptic site, and thus blocking fusion and clustering of SNARE-associated vesicles to the synaptic PM [152]. This would lead to neurotransmitter release inhibition, increased synaptic dysfunction and neuronal cell death.

1.4.2.5 α -syn aggregation alters modulation of P-body components

In the nucleus, N-terminus of α -syn interacts closely with a core set of proteins (enhancer of messenger RNA (mRNA)-decapping 3 (Edc3), Edc4, mRNA-decapping enzyme subunit 1 (Dcp1), Dcp2, and 5'-3' exoribonuclease 1 (Xrn1)) responsible for modulating processing body (P-body) complexes in the cytosol [153]. P-bodies are crucial for the regulation of mRNA stability, mRNA storage, mRNA translation, mRNA decapping and degradation. P-bodies also contribute to protein quality control processes through RNA-binding proteins (RBPs) [154]. As α -syn pathologically accumulates, aberrant interaction with Edc4 occurs at the expense of physiological decapping-

module interactions. mRNA decay kinetics within PD-relevant pathways are disrupted in PD patient neurons and the brain [153, 154].

1.5 Summary

Understanding the complexities of Alzheimer's disease (AD) and Parkinson's disease (PD), the most prevalent neurodegenerative disorders (NDs), is essential for improved diagnosis and treatment. These diseases are characterized by the progressive loss of neuronal function and share proteinopathy (like A β oligomers and p-tau in AD and aggregated α -syn in PD) as a common pathological hallmark. Furthermore, both are associated with aberrant protein clearance mechanism dysfunctions. Other cellular mechanisms, including mitochondrial dysfunction, oxidative stress, impaired lysosomal pathways, and neuro-inflammatory reactions triggered by glial cell activation, also contribute to the progression of these multifaceted disorders. While these NDs share several biological events, they differ in risk factors, cellular types affected, and the genes implicated.

In this context, the objectives of this work were threefold:

- 1- To discover targets associated with specific traits in AD and PD.
- 2- To identify common signatures and shared mechanisms between AD and PD.
- 3- To investigate whether central disease-related signatures can be discerned in the peripheral blood of patients with these conditions.

Corresponding to these objectives, I proposed the following hypotheses:

- 1- The transcriptomic profiles from the central brain (DLPFC/BA9) and peripheral blood (PBMC and WB) of AD and PD patients will reveal differentially expressed features unique to each disease.
- 2- While AD and PD have their distinctions, their transcriptomic profiles will highlight shared signatures and mechanisms.
- 3- Central disease-related transcriptomic signatures will be detectable in peripheral blood samples from AD and PD patients.

To validate these hypotheses, I analysed human RNA-seq data from post-mortem central brains (DLPFC/BA9) and live patient peripheral blood samples (PBMC and WB) of those diagnosed with AD and PD. Utilizing production bioinformatics, high-performance computing, and streamlined tools, I aimed to uncover features with potential clinical diagnostic value for both diseases.

Chapter 2 focuses on selecting pivotal features (gene/transcriptome) using differential gene expression combined with a random forest-based machine learning algorithm. Chapter 3 delineates the process of RNA-seq variant (SNP) calling, emphasizing the association of genomic variations with the disease. In Chapter 4, a combination of expression and splicing quantitative trait loci (eQTL and sQTL) mapping provides further insights into the role of markers associated with AD and PD risks. Lastly, Chapter 5 synthesizes the findings from the three studies, presenting overarching conclusions.

Chapter 2 Overlap between central and peripheral transcriptomes in Parkinson's disease but not Alzheimer's disease

The sections, figures, tables of Chapter 2 have been taken from my published manuscript.

Overlap between Central and Peripheral Transcriptomes in Parkinson's Disease but Not Alzheimer's Disease

Kosar Hooshmand, Glenda M. Halliday, Sandy S. Pineda, Greg T. Sutherland and Boris Guennewig

Published in: *International Journal of Molecular Sciences*.

Author contributions

The undersigned certified that the following is the fair and accurate description of the individual contributions made by each other to this publication:

Kosar Hooshmand,

Wrote the manuscript & prepared figures and tables

Glenda M. Halliday,

Co-wrote the manuscript

Sandy S. Pineda,

Co-wrote the manuscript

Greg T. Sutherland,

Co-wrote the manuscript

Boris Guennewig,

Co-wrote the manuscript

2.1 Introduction

Alzheimer's disease (AD) and Parkinson's disease (PD) are well-known neurodegenerative disorders (NDs) that share common pathological events. They correlate with changes in gene expression that occur before the onset of and during the progression of these diseases [155]. Brodmann Area 9 (BA9) of the dorsolateral prefrontal cortex (DLPFC) plays an essential role in cognitive, motor, and memory-related functions and is pathologically affected in both diseases. AD and PD both feature prion-like spread of pathology [155] such that at postmortem, in any particular case, there will be brain regions at different stages of the disease [139, 140]. Notably, at postmortem, the BA9 has mild to moderate pathological changes with most neurons intact. In this sense, analysis of BA9 transcriptome profiles may provide insights related to early neurodegenerative mechanisms involved in AD and PD [155, 156].

Given the relative inaccessibility of the brain during a person's lifetime, peripheral biomarkers that accurately reflect the state and progression of brain diseases and the responses to potential treatments are regarded as the 'holy grail' in the clinical management of NDs. Such markers have remained particularly elusive, although there is current optimism for a phosphorylated form of tau (p-tau217) as a blood biomarker in AD [37]. Blood biomarkers associated with central neuronal system (CNS) pathology could reflect the influence of the same genetic variants peripherally, systemic effects of the disease processes, and despite the bidirectional nature of the blood-brain-barrier (BBB), leakage or active transport of brain-borne pathological entities. Specifically, the release of mRNAs from apoptotic neurons into the plasma has already been described [37, 157]. Given the common pathogenic mechanisms in NDs within the CNS, it is possible that these might also be observed peripherally.

Supervised machine learning algorithms show great promise for the analysis of 'big data' in ND research towards the diagnosis, prognosis, and development of new therapies [17]. Random forest,

a supervised machine learning approach, has shown essential advantages over other methodologies, including the ability to handle highly nonlinearly correlated data, robustness to noise, tuning simplicity, and achieving the highest ratio of self-consistent selections in its results [158, 159]. Therefore, the current study specifically applied a random forest-based algorithm with the feature selection tool, Boruta [160], to explore the transcriptome of the BA9, WB, and PBMC in human AD, PD, and cognitively healthy controls (NC) using RNA-sequencing (RNA-seq) data held in the National Centre of Biotechnological Information's sequence read archive (SRA; <https://www.ncbi.nlm.nih.gov/sra> (accessed on 1 September 2018)).

In theory, large data dimensions add more information to the dataset, thereby improving the quality of the data. However, one of the biggest challenges with the majority of public datasets, including SRA, is that there are more features or transcripts = p observed than patients = n characterised. This problem is known as the “curse of dimensionality” and substantially impacts analysis accuracy and recall due to overfitting that can mistake small fluctuations for significant variance in the data, and lead to spurious findings [18]. Therefore, the current study accessed the largest sample sizes held in SRA and utilised production bioinformatics, high-performance computing, and streamlined tools, coupled with differential gene expression and machine learning approaches to reduce noises as much as possible to avoid unnecessary complexity in the inferred models and improve the algorithm's efficiency. The findings allowed for a comprehensive view of the publicly available RNA-seq data in AD/PD and provide some evidence of shared genes and molecular pathways driving NDs, but primarily show distinct disease-specific patho-mechanisms both within the brain and peripherally for these NDs.

2.2 Materials and methods

2.2.1 Data collection

To determine predictive NDs genetic variants and their biological function, the ERA, Array Express and ENCODE databases along with the sequence read archive (SRA) (<https://www.ncbi.nlm.nih.gov/sra> (accessed on 1 September 2018)) database (Figure 1A) were searched to collate all publicly available RNA sequencing (RNA-seq) studies derived from (i) human postmortem brain tissue and (ii) peripheral blood from patients with AD, PD, and healthy controls. The keywords utilised in the initial search were as follows: ‘Parkinson’s disease AND Prefrontal Cortex’, ‘Parkinson’s disease AND Blood’, ‘Alzheimer’s disease AND Prefrontal Cortex’, and ‘Alzheimer’s disease AND Blood’.

The initial search yielded 835 SRA fastq files, which recorded a total of 26 studies. Subsequently, all selected fastq files were downloaded using the `getFASTQfile` function from the `SRADB` R/Bioconductor package [161]. All studies downloaded encompassed data from postmortem brain region: Brodmann area 9 only (BA9) (n = 155 with AD, n = 172 with PD, and n = 184 NC), and blood: PBMCs datasets (n = 33 with AD, n = 47 with PD, and n = 78 NC) and WB (n = 54 with AD, n = 38 with PD, and n = 74 NC) summarised in Table 2.1. Metadata obtained from each study were used to classify patients into NC, AD, and PD groups. In the SRA dataset, some of the AD and PD patients also had a definitive diagnosis according to Braak stage values, RNA integrity number (RIN), age of death, and gender (Appendix 1).

2.2.2 Workflow implementation and data processing

All steps of the pipeline were executed using a workflow designed in the Workflow Description Language (WDL). This workflow utilized Docker images, operating in Cromwell, as illustrated in Figure 2.1 B. Detailed scripts and workflow steps are available in <https://github.com/kohsh/SRA-RNAseq>.

For the preprocessing of raw sequences, Trimmomatic (v0.38) [162] was employed to perform adapter removal, quality trimming, and to filter reads based on a minimum length requirement using default settings. It's worth noting that while the use of default Trimmomatic parameters may not be optimal for every sample, our downstream sample grouping, selection, and filtering strategies are designed to counteract potential biases or technical artifacts introduced during preprocessing. These rigorous strategies ensure the robustness of the findings despite potential variations in cDNA synthesis protocols, library preparation strategies, and sequencing technologies across the datasets that were analysed.

Post-preprocessing, FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) was utilised to generate a quality report for each processed read file (https://github.com/binfstats/SRA-RNAseq/blob/main/Scripts/Trim-%26_QC/Trim-QC.wdl). The trimmed RNA-seq reads were subsequently aligned against the human genome version 38 (GRCh38, gencode.v31.transcripts.fa) using the pseudoaligner Kallisto [163] with default settings (<https://github.com/binfstats/SRA-RNAseq/blob/main/Scripts/Kallisto/Kallisto.wdl>). Kallisto was used as it strikes a balance between mapping runtime and similar or better accuracies compared with other alignment methods [164].

2.2.3 Transcriptomic data normalisation and expression analysis

Genes with low expression in all samples were removed using the “filterByExpr” function in the edgeR R/Bioconductor package [165], leaving 39,500 genes for downstream analysis. To correct any compositional bias from our data, the Trimmed Mean of M-values (TMM) normalization method [166] was used from the “calcNormFactors” function in the edgeR package. Trimmed samples from different studies were randomly subdivided into groups using the “sample” function in R.

To account for potential batch effects arising from differences in experimental conditions, sample processing, or other technical variables in the transcriptomic data, batch effect removal was performed. Following normalisation, the “removeBatchEffect” function from the limma R/Bioconductor package was used. Specifically, the “removeBatchEffect” function employs a linear model to estimate and remove the unwanted variation associated with the identified batches. In the analysis, the batch variable (e.g., sample collection date, sequencing run) in the function was specified to adjust the data for these effects. After batch effect removal, the adjusted expression matrix was used for downstream analysis.

Principal components analysis, hierarchical clustering, and ANOVA test were performed and visualized using the R/Bioconductor package. Differential expression analyses were conducted in a negative binomial model-based method using edgeR. Differentially expressed (DE) genes were filtered for those having a gene expression level Benjamini–Hochberg (BH) Q value < 0.05 , and minimum fold change $|\log_{10}FC| \geq 1$, with false-discovery rate (FDR) adjusted p-values ≤ 0.05 .

2.2.4 Analysis of genes associated with NDs by machine learning (ML)

One of the biggest problems in a differential expression (DE) analysis of RNA-seq expression data is their high dimensionality and high pairwise correlations between genes. Genes may exhibit high correlation if they participate in the same biological pathway or are co-regulated [15, 167]. Such high correlation can pose challenges in differential expression analysis for a couple of primary reasons. First, it can introduce redundancy, as genes that are highly correlated often offer similar insights into the biological process under scrutiny. This redundancy might overemphasize the significance of certain genes, thereby shadowing the identification of truly pivotal ones. Second, in regression-based methodologies, high correlation can induce multicollinearity, leading to unstable coefficient estimates and complicating the interpretation of individual predictors [15,

167]. Given the complications that arise from high correlations among DEGs, using all DEGs becomes counterproductive when developing diagnostic and prognostic prediction tools. To address these problems and to further identify all relevant features in the datasets, the Boruta algorithm, a wrapper that is built around the random forest (RF) classifier [168, 169] was set to be used. Boruta-ML was run with max runs set to 20,000 for both classification and regression. The features which have importance significantly higher than the maximum Z score (MZS) were treated as relevant or confirmed.

2.2.5 Statistical assessment of overlapping genes

The hypergeometric test was utilised to assess the significance of overlapping genes derived from different conditions and tissues. Given the entire number of analysed genes, this test is adept at determining the likelihood of observing a specific overlap between two gene sets.

The parameters for the hypergeometric test were defined as:

N : The total number of genes in the human genome that were considered in our study.

K : Total genes pinpointed in the condition (1) /tissue (1).

n : Total genes identified in the condition (2) /tissue (2).

k : The observed overlap or the number of genes found in common between condition (1) /tissue (1) and condition (2) /tissue (2).

Given these parameters, the probability $P(X \geq k)$ of observing an overlap of or more genes by random chance were computed using the hypergeometric distribution.

2.2.6 Gene ontology and pathway enrichment analysis

To learn the potential functions of significant genes, topGO, an R/Bioconductor package [170], was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genome

(KEGG) enrichment analysis. Top-ranked GO biological pathways with an adjusted p-value of <0.05 were identified and used to select overrepresented or most enriched genes.

2.2.7 Gene–gene interaction and gene co-expression network analysis

GeneMANIA webserver was used to identify genetic interaction (GI) and to represent co-expressed genes that are functionally coordinated with the selected gene biomarkers in response to a similar condition [171]. The study supports the better prediction of the most important functional genes that might provide a more robust bio-signature for phenotypic traits, thus providing more suitable biomarker candidates for future studies.

2.3 Results

2.3.1 Non-overlapping sample subset selection using differentially expressed genes

Aiming to uncover the molecular reconfigurations underlying PD and AD, the two most common NDs, large-scale RNA-seq studies were accessed through the Sequence Read Archive (SRA) database (Table 2.1, Figure 2.1 A and Appendix 1). Due to variations in specific study designs, a subset of non-overlapping samples from the BA9, whole blood (WB), and PBMC in human AD, PD, and cognitively healthy controls was selected (Table 2.1 and Appendix 2). During our data preprocessing phase, Trimmomatic was employed to trim and quality-filter the raw reads sourced from the SRA. Post-trimming, the processed reads were aligned to the GRCh38 reference genome, specifically the gencode.v31.transcripts.fa, using a comprehensive bioinformatics pipeline (Figure 2.1 B, and <https://github.com/binfstats/SRA-RNAseq>).

Table 2.1 Number of tissue-specific samples per phenotype.

Condition	PBMC	WB	DLPFC/BA9	PBMC	WB	DLPFC/BA9
	Total number of samples			Final number of samples		
Parkinson's Disease	47	38	172	6	20	126
Alzheimer's Disease	33	54	155	22	48	101
Cognitively Healthy Controls	78	74	184	25	42	162

After processing, the data from SRA, categorized by tissue type and specific diseases, were merged and normalized, as visualized in Figure 2.1 C. Then principal components (PCA) was computed for all samples represented in Figure 2.2 and proceeded with hierarchical clustering (illustrated in Figure 2.3). During PCA analysis, the potential for batch effects were recognized, which could arise from variations in library preparation and sequencing methods across different studies. To mitigate their impact, samples were grouped based on metadata factors such as studies (project-IDs), and strategy (sequencing platform). Additionally, by selecting samples representing human AD and PD conditions, the influence of inter-individual differences was reduced. Our downstream analysis pipeline further helped identify consistent biological signals associated with these diseases.

Subsequently, an ANOVA test was performed, considering several confounding covariates/factors that might impact our study's results. These factors include Condition, which represents the distinct groups under comparison, thereby inherently carrying possible biases. RIN was included due to its influence on gene expression outcomes, especially considering that degraded RNA in samples with low RIN could potentially skew results. The Braak stage was incorporated as an indicator of disease progression, which might impact gene expression patterns as the condition advances. Age of Death was factored in to address age-associated variations in gene expression. Sex was included to cater to biological differences in gene expression between genders. Study (project-IDs) was also accounted to control potential variations arising from differing methodologies across research initiatives. Lastly, Strategy was considered to handle any biases or discrepancies that could emerge from different sequencing methodologies. These confounders were systematically integrated into our analysis to enhance our study's robustness, as illustrated in Figure 2.4. Differentially expressed genes (DEGs) were identified for all matched independent subtypes by comparing the subtypes to the other samples using the negative binomial distribution model, and the exact test was performed

(Table 2.2). Functional enrichment of the DEGs was also tested to identify GO terms associated with each of the subtypes (Table 2.3). Accordingly, combined independent studies with GO terms that showed clear relation to the disease/tissue of interest were established as the final selected subsets.

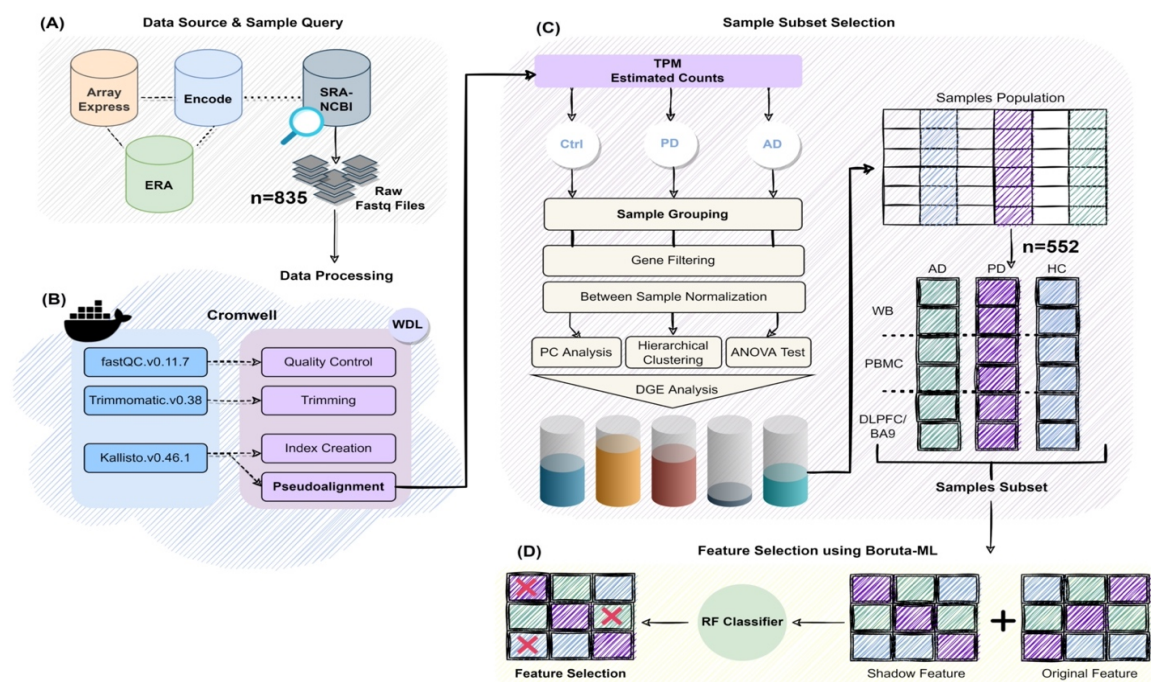


Figure 2.1 Schematic representation showing all steps necessary for RNA-seq analysis.

(A) The Sequence Read Archive (SRA) database that makes an effort to collect the publicly available transcriptomics held in ERA, Array Express and ENCODE databases was searched for RNA-seq data to access publicly available human postmortem brain- and blood-based studies. (B) All the processing steps of the pipeline were wrapped into WDL tasks that were designed to be executed on the cloud-based services with Cromwell. Tasks in WDL workflow have an associated Docker-based tool image since WDL does not directly have the concept to build a tool. (C) Summary of the steps performed for the selection of a subset of samples using differential expression meta-analysis paradigm. (D) Boruta random forest (RF)-based algorithm was used as a feature selection method on normalised data. Boruta adds randomness to the given dataset by creating shuffled copies of all features, which are called shadow features. Then, it trains a random forest classifier on this extended data and applies a feature importance measure and evaluates the importance of each feature.

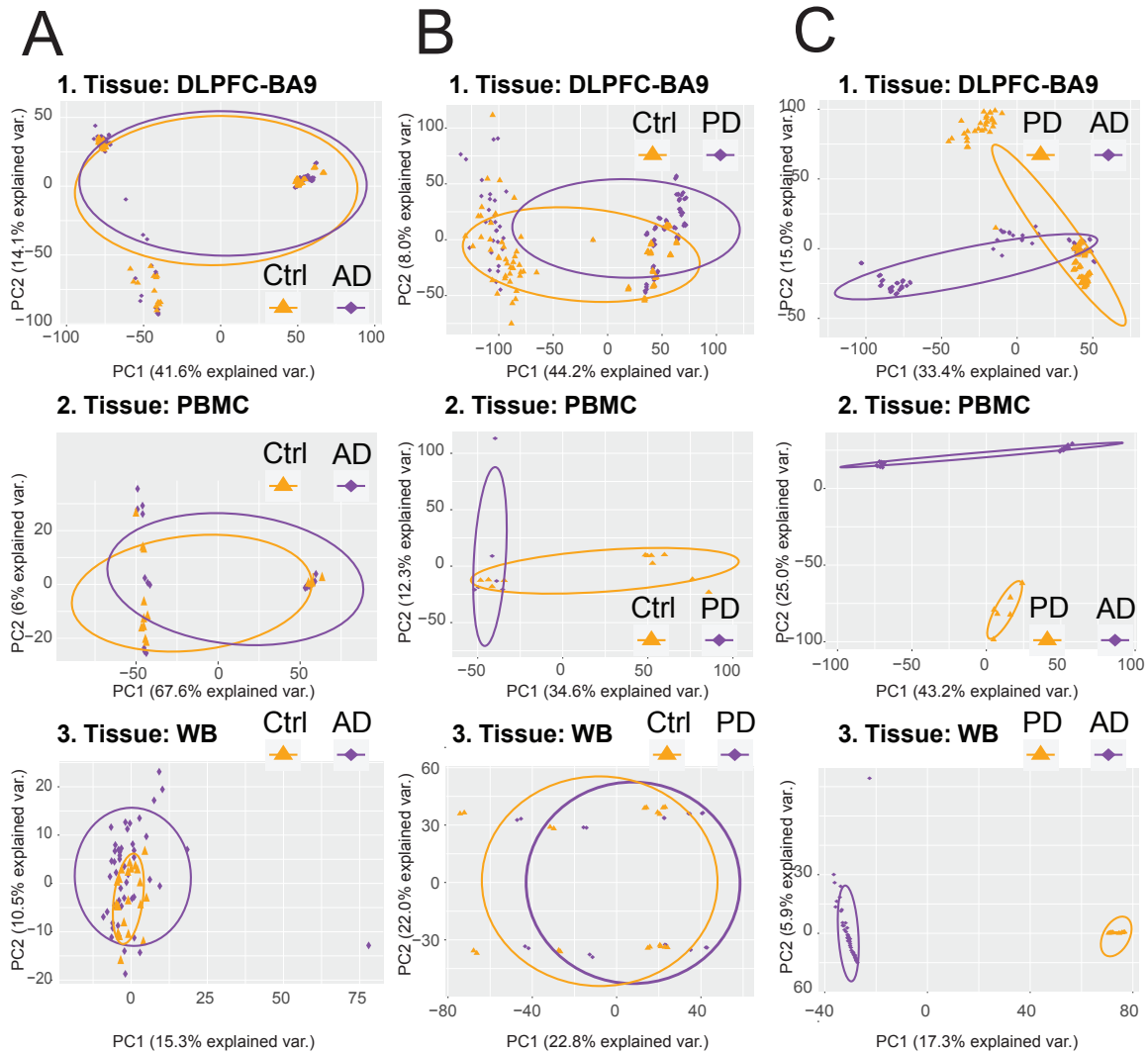


Figure 2.2 Score plots for the first two Principal Components (PCs) of comparing groups.

Graphs represent the separation of the sample populations of comparing groups **(A)** AD-vs.-ctrl, **(B)** PD-vs.-ctrl, and **(C)** AD-vs.-PD) in different tissues **(1)** DLPFC/BA9, **(2)** PBMC, and **(3)** WB into clusters based on the first PC (PC1).

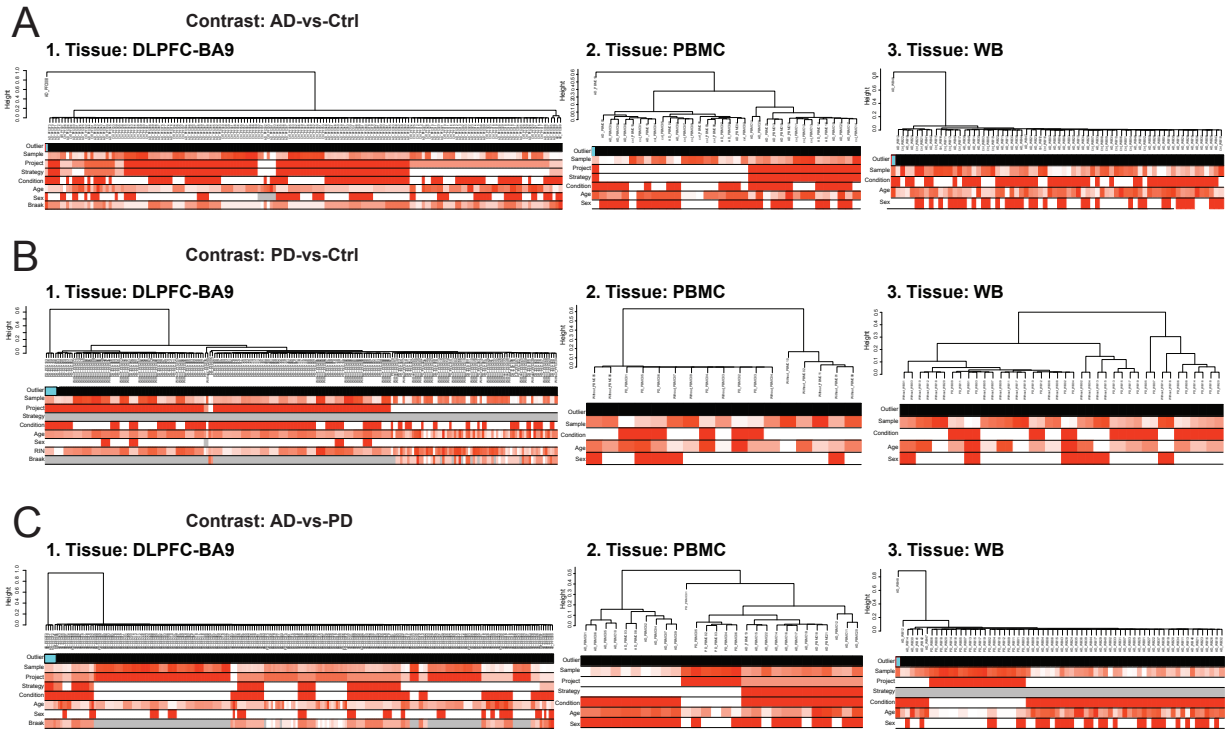


Figure 2.3 Hierarchical clustering for different tissue in comparing groups.

Cluster trees for **(A)** AD-vs.-ctrl, **(B)** PD-vs.-ctrl, and **(C)** AD-vs.-PD in different tissues **(1)** DLPFC-BA9, **(2)** PBMC, and **(3)** WB. The leaves of the tree correspond to the samples in each contrast. The first color band underneath the tree indicates which sample network appear to be outlying (Light blue box) compared to the ones that are grouped (solid black box). The remaining color-bands represent numerical values of physiologic traits (Red indicates high values).

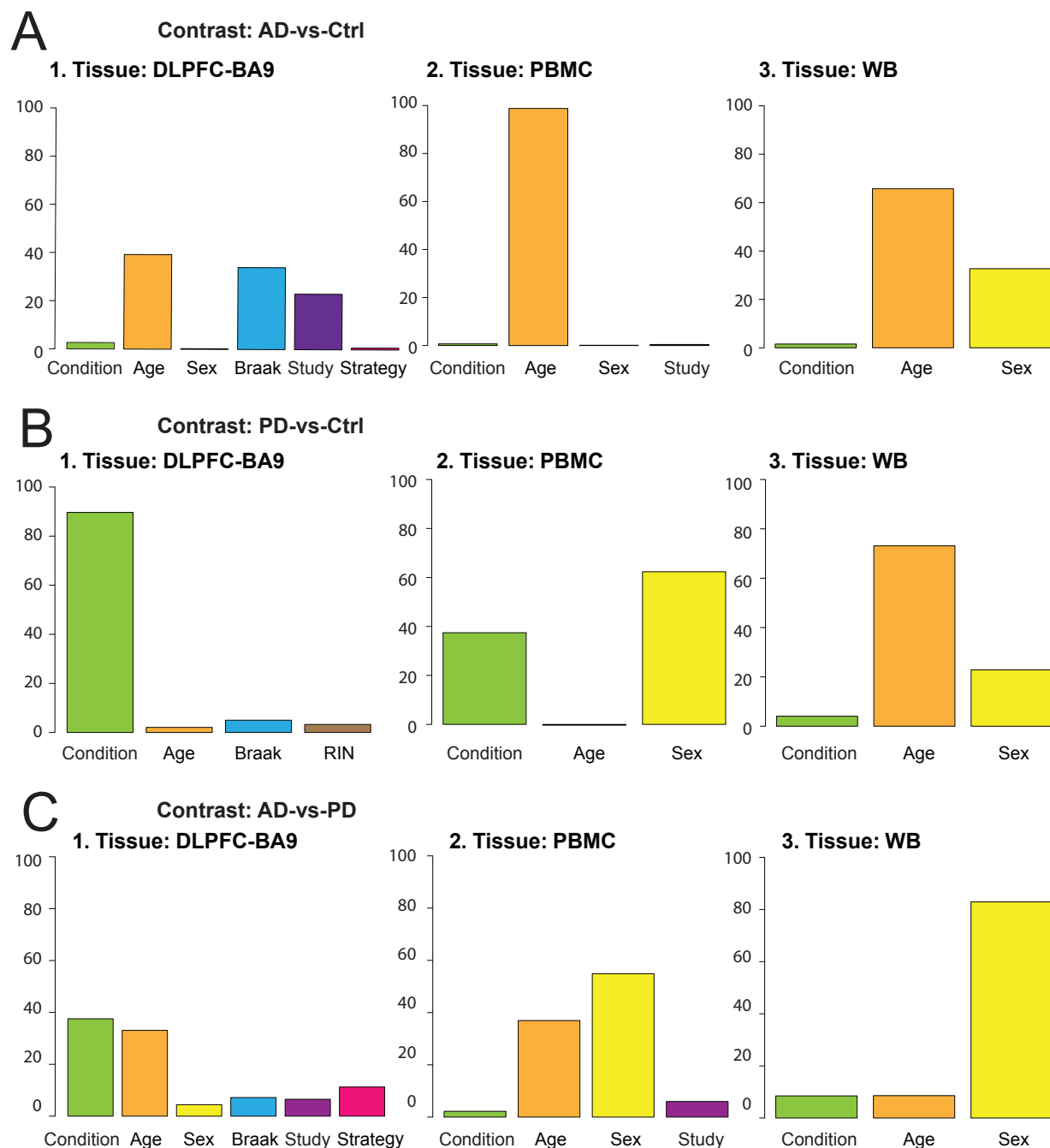


Figure 2.4 Schematic representation of ANOVA test for comparing groups.

Schematic representation of ANOVA test results used to research the effects of multiple factors (Condition: describes the comparing groups, Age, Sex, Braak, Study (project-IDs): describes the goal of the research effort, and Strategy: describes the sequencing strategy) on expression data from the sample populations of comparing groups (A) AD-vs.-

ctrl, (B) PD-vs.-ctrl, and (C) AD-vs.-PD) in different tissues (1) DLPFC-BA9, (2) PBMC, and (3) WB. Y axis represents the percentage (%) of effectiveness.

Table 2.2 Top ten differentially expressed genes (DEGs) for comparing groups.

Contrast	Tissue	Chr	Start	End	Strand	Gene	logFC	P-value	FDR
AD_vs._ctrl	DLPFC/ BA9	20	37693955	37872129	+	CTNNBL1	2.024819	6.30E-14	2.83E-10
		16	176680	177522	+	HBA1	7.815985	6.30E-14	2.83E-10
		18	21854640	21854987	-	-	-2.87622	5.05E-13	1.51E-09
		21	8759077	8761335	+	LINC01666	-3.35782	1.79E-12	3.72E-09
		19	49487510	49493057	+	RPL13A	6.79068	2.48E-12	3.72E-09
		22	26646411	26676475	+	MIAT	-2.15216	2.49E-12	3.72E-09
		20	35117136	35117207	-	-	-3.11073	6.11E-12	7.85E-09
		22	24594811	24629005	+	GGT1	3.227733	2.29E-11	2.57E-08
		12	6873569	6914241	+	LRRC23	1.726165	4.65E-11	4.64E-08
		1	15941869	15976132	-	ZBTB17	1.817886	5.55E-11	4.98E-08
PD_vs._ctrl		1	2.6E+07	2.6E+07	-	STMN1	-2.76189	9.20E-25	1.11E-20
		1	631074	632616	+	MTCO1P12	2.17243	6.90E-20	4.15E-16
		12	4.6E+07	4.6E+07	-	SLC38A2	-0.94787	2.66E-19	1.07E-15
		14	7.6E+07	7.6E+07	-	TGFB3	-1.21138	6.95E-18	1.39E-14
		17	4.9E+07	4.9E+07	+	CALCOCO2	-1.62465	2.05E-16	3.53E-13
		7	7.4E+07	7.4E+07	-	STX1A	1.0379	7.91E-12	1.36E-09
		17	8159149	8163546	-	VAMP2	0.5729	1.51E-11	2.02E-09
AD_vs._PD		17	7.9E+07	7.9E+07	+	SYT1	-1.04602	0.00036	0.0016
		17	6.9E+07	6.9E+07	+	PRKAR1A	-16.5975	9.17E-308	1.29E-305
		1	1.7E+08	1.7E+08	+	DNM3	-21.3113	9.98E-306	1.38E-303
		17	1344275	1400222	-	YWHAE	22.3746	2.09E-303	2.85E-301
		15	4.6E+07	4.6E+07	+	-	-14.7842	3.43E-301	4.60E-299
		4	1.2E+08	1.2E+08	+	KIAA1109	-18.1197	5.26E-301	6.97E-299
		22	3.2E+07	3.2E+07	+	YWHAH	-22.1041	2.08E-300	2.72E-298
		6	5.6E+07	5.7E+07	-	DST	-18.8116	2.78E-297	3.56E-295
		X	5.4E+07	5.4E+07	-	HUWE1	-12.2943	3.28E-295	4.15E-293
		8	1.1E+08	1.1E+08	+	OXR1	-17.9359	2.09E-293	2.61E-291
AD_vs._ctrl		6	4706159	4955551	+	CDYL	-21.5392	5.17E-293	6.35E-291
		14	7.4E+07	7.4E+07	-	MIDEAS	-4.39633	5.89E-05	0.09414
		3	4.6E+07	4.6E+07	+	CCR2	-4.2794	6.31E-05	0.09414
	15	4.9E+07	4.9E+07	+	EID1	-2.61101	7.20E-05	0.09414	
	3	1.6E+08	1.6E+08	-	KPNA4	3.88839	0.00017	0.13698	
	16	3.4E+07	3.4E+07	-	-	-3.09502	0.00021	0.13698	
	14	1.1E+08	1.1E+08	-	IGHJ6	-3.51848	0.00023	0.13698	
	X	1.5E+08	1.5E+08	+	GDI1	-3.25799	0.00024	0.13698	
	2	1.6E+08	1.6E+08	+	-	-3.65452	0.00033	0.16377	
	MT	14149	14673	-	MT-ND6	-2.59214	0.00046	0.20197	
PD_vs._ctrl	14	1.1E+08	1.1E+08	-	IGHV5-10-1	-4.88888	0.00055	0.21676	
	14	22536145	22536204	+	TRAJ8	9.2154169	4.9907406	1.30E-08	
	1	16514122	16514285	-	RNU1-1	12.127531	12.808968	2.32E-06	
	1	16666785	16666948	-	RNU1-3	12.127531	12.808968	2.32E-06	
	1	16740516	16740679	+	RNU1-4	12.127531	12.808968	2.32E-06	
	1	16895980	16896143	+	RNU1-2	12.127531	12.808968	2.32E-06	
	1	143729407	143729570	-	RNVU1-18	12.127531	12.808968	2.32E-06	
	1	146376807	146376970	+	RNVU1-29	12.127531	12.808968	2.32E-06	
	14	34546714	34546877	+	RNU1-27P	12.127531	12.808968	2.32E-06	
	14	34556226	34556389	-	RNU1-28P	12.127531	12.808968	2.32E-06	
AD_vs._PD	3	24521362	24521435	-	-	18.395584	14.210279	5.57E-06	
	14	49862550	49862849	-	RN7SL2	36.19159	1.20E-215	8.76E-212	
	19	8520778	8577577	-	MYO1F	-22.35438	1.00E-132	3.65E-129	
	11	62559596	62574086	-	EEF1G	21.69292	1.92E-124	4.68E-121	
	6	26156329	26157115	+	H1-4	23.71923	1.22E-117	1.87E-114	
	3	49357176	49358353	-	GPX1	20.12463	1.46E-117	1.87E-114	
	12	56116590	56117967	+	RPL41	27.34022	1.54E-117	1.87E-114	
	17	74203582	74210655	+	RPL38	22.66484	3.17E-117	3.30E-114	
	14	1.06E+08	1.06E+08	-	IGHV4-34	-24.95384	1.06E-114	9.68E-112	
	15	72199029	72231822	-	PKM	21.26962	2.90E-112	2.35E-109	
AD_vs._ctrl	WB	12	49707725	49764934	+	TMBIM6	19.94127	5.36E-112	3.91E-109
		7	6.6E+07	6.6E+07	+	TPST1	3.37013	1.06E-10	7.16E-08
		1	5.8E+07	5.8E+07	+	DAB1-AS1	-3.8556	1.94E-10	7.16E-08
		17	8.3E+07	8.3E+07	+	-	-3.8752	6.59E-10	1.62E-07

PD_vs._ctrl	7	1.6E+08	1.6E+08	+	-	-3.6901	4.69E-09	8.67E-07
	16	6.7E+07	6.7E+07	+	-	1.45361	8.30E-08	1.23E-05
	5	1.1E+08	1.1E+08	+	TSLP	1.92572	1.09E-07	1.35E-05
	X	1.9E+07	1.9E+07	-	PHKA2	-2.8616	2.11E-07	2.23E-05
	19	1.7E+07	1.7E+07	-	-	1.64634	3.65E-07	3.37E-05
	2	2E+08	2E+08	-	SF3B1	1.58627	4.84E-07	3.96E-05
	17	7.6E+07	7.6E+07	-	CASKIN2	1.88522	5.35E-07	3.96E-05
	9	1.28E+08	1.28E+08	+	LCN2	1.428239	6.79E-14	3.83E-10
	8	6935820	6938306	-	DEFA4	1.767847	1.76E-11	4.96E-08
	8	6977649	6980092	-	DEFA1	1.968336	4.27E-10	7.64E-07
	14	1.06E+08	1.06E+08	-	IGHG4	-2.12341	9.28E-08	9.72E-05
	3	48223347	48225491	+	CAMP	1.032794	3.23E-07	0.000228
	14	24606480	24609699	-	GZMH	-0.8962	1.24E-06	0.000623
	17	8159149	8163546	-	VAMP2	-0.06097	0.608882	0.917201
	11	62806860	62832051	-	STX5	-0.29218	0.092437	0.59302
	6	1.32E+08	1.33E+08	-	STX7	0.327919	0.118417	0.64493
	6	1.44E+08	1.44E+08	+	STX11	-0.26687	0.104108	0.615884
	1	27342020	27353937	+	SYTL1	-0.16051	0.35069	0.840469
AD_vs._PD	2	88857161	88857683	-	IGKC	16.88149	7.18E-270	2.26E-266
	6	31268749	31272130	-	HLA-C	15.46771	2.78E-253	4.36E-250
	7	5526409	5563902	-	ACTB	14.54015	1.03E-251	1.08E-248
	6	30489509	30494194	+	HLA-E	13.43758	8.97E-247	7.05E-244
	5	1.5E+08	1.5E+08	-	CD74	15.07285	1.25E-217	7.86E-215
	22	39312882	39320389	-	RPL3	13.52419	3.06E-217	1.60E-214
	12	14942031	14961728	-	ARHGDIB	14.37452	4.77E-217	2.14E-214
	10	71816298	71851325	-	PSAP	13.61667	4.88E-211	1.92E-208
	6	31353872	31357188	-	HLA-B	14.62045	1.11E-210	3.87E-208
	6	32439878	32445046	+	HLA-DRA	13.71302	4.81E-210	1.51E-207

Abbreviation: Chr: chromosome, logFC: log-fold change threshold (-logFC = downregulation of genes, +logFC = upregulation of genes), FDR: false discovery rate.

Table 2.3 Number of differentially expressed genes (DEGs) and enriched GO-BPs from a selected subset of samples.

Tissue	Contrast	Number of DEGs	Top GO-BPs	Genes	P-value
DLPFC/ BA9	AD-vs.-ctrl	8948	1-Intracellular transport 2-Cellular component organisation 3-Cellular protein localisation 4-Organelle organisation 5-Protein localisation 6-mRNA metabolic process 7-Peptide transport 8-Nitrogen compound transport	Chondroitin sulfate proteoglycan 5 (CSPG5), DAAMI, SEPTIN9, kinesin family member 5C (KIF5C), Unc-51 like autophagy activating kinase 1 (ULK1), Ubiquitin-specific protease 9, X-LINKED (Usp9X), RAP2A, Transforming growth factor beta regulator 4 (TBRG4), TAP binding protein (TAPBP), Solute carrier family 6 member 8 (SLC6A8)	1-3.48E-98 2-1.16E-95 3-3.68E-87 4-6.66E-85 5-1.42E-80 6-7.28E-70 7-2.14E-66 8-6.47E-60
	PD-vs.-ctrl	12,043	1-Intracellular transport 2-Cellular protein localization 3-Cellular component organisation 4-Protein localisation 5-Intracellular protein transport 6-Cellular protein metabolic process 7-Catabolic process 8-mRNA metabolic process	RANBP1, Spire type actin nucleation factor 1 (SPIRE1), Solute carrier family 9 member A3 (SLC9A3), SEPTIN9, Peroxisomal biogenesis factor 10 (PEX10), VTI1B, Transmembrane protein 132A (TMEM132A), Trans-golgi network protein 2 (TGNL2), Phosphotriesterase-related protein (PTER), Poly (RC) binding protein 2 (PCBP2)	1-1.76E-160 2-5.02E-145 3-5.98E-132 4-7.78E-127 5-2.17E-112 6-2.15E-89 7-7.42E-75 8-7.52E-73
	AD-vs.-PD	8554	1-Intracellular transport 2-Cellular protein localisation 3-Protein localisation 4-Cellular component organisation 5-Intracellular protein transport 6-Organelle organisation 7-Peptide transport 8-Nitrogen compound transport	VPS41, SEPTIN9, ULK1, dishevelled associated activator of morphogenesis 1 (DAAM1), Dishevelled associated activator of morphogenesis 2 (DAAM2), Transmembrane P24 Trafficking Protein 7 (TMED7), Golgi reassembly-stacking protein 2 (GORASP2), TAPBP, SLC6A8	1-7.83E-124 2-1.85E-108 3-3.41E-100 4-1.73E-99 5-1.81E-83 6-3.32E-79 7-2.03E-78 8-8.40E-74
WB	AD-vs.-ctrl	740	1-SRP-dependent co-translational protein targeting to membrane 2-Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay 3-Viral process 4-Viral transcription 5-Nuclear-transcribed mRNA catabolic process 6-Protein localisation to membrane	Ribosomal protein L31 (RPL31), Ribosomal protein L32 (RPL32), SMG5, H2AX, LSM4, Adaptor related protein complex 3 subunit delta 1 (AP3D1)	1-1.54E-08 2-9.60E-08 3-5.27E-07 4-6.46E-07 5-5.58E-06 6-6.36E-06
	PD-vs.-ctrl	5641	1-Granulocyte activation 2-Neutrophil degranulation 3-Neutrophil activation involved in immune response 4-Leukocyte degranulation 5-Neutrophil activation 6-Neutrophil mediated immunity	Vesicle associated membrane protein 8 (VAMP8), Myeloid differentiation primary response 88 (MYD88), Spleen associated tyrosine kinase (SYK), HCK, C-X-C Motif chemokine receptor 2 (CXCR2), WD repeat domain 1 (WDR1), Fc alpha receptor (FCAR), TYROBP, SYK	1-2.81E-134 2-5.64E-133 3-6.60E-133 4-7.40E-133 5-2.31E-132 6-5.35E-131 7-4.88E-126

			7-Myeloid leukocyte activation 8-Leukocyte activation involved in immune response		8-4.72E-119
	AD-vs.-PD	3143	1-Neutrophil activation 2-Neutrophil activation involved in immune response 3-Neutrophil degranulation 4-Neutrophil mediated immunity 5-Myeloid leukocyte activation 6-Leukocyte activation 7-Regulated exocytosis 8-Vesicle-mediated transport	CXCR2, C-C motif chemokine ligand 5 (CCL5), Fc epsilon receptor Ig (FCER1G), TYRO protein tyrosine kinase binding protein (TYROBP), Stimulator of interferon response CGAMP interactor 1 (STING1), Major histocompatibility complex, class I, B (HLA-B), Major histocompatibility complex, class I, C (HLA-C), WDR-1, Peroxiredoxin 1 (PRDX1), PRDX2, Synaptogyrin 2 (SYNGR2), Myosin heavy chain 9 (MYH9), Reticulon 3 (RTN3), COPI coat complex subunit gamma 1 (COPG1), Perilipin 3 (PLIN3), ERGIC and Golgi 3 (ERGIC3)	1-1.46E-114 2-3.91E-114 3-6.57E-114 4-2.38E-111 5-2.30E-107 6-4.45E-99 7-4.48E-96 8-3.15E-81
PBMC	AD-vs.-ctrl	3921	1-mRNA metabolic process 2-Intracellular transport 3-Cellular protein localisation 4-Translational initiation 5-Cellular metabolic processes 6-Cotranslational protein targeting to membrane 7-Protein targeting to ER 8-SRP-dependent cotranslational protein targeting to membrane	Poly (RC) binding protein 2 (PCBP2), RNA binding protein 1 (RNABP1), SEPTIN9, ATP binding cassette subfamily E member 1 (ABCE1), Exosome component 10 (EXOSC10), SEC61 translocon subunit alpha 1 (SEC61A1), Translocation associated membrane protein 1 (TRAM1), Signal recognition particle 14 (SRP14), SRP receptor subunit alpha (SRPRA), Ribosomal protein L31 (RPL31), Ubiquitin A-52 residue ribosomal protein fusion product 1 (UBA52)	1-1.18E-84 2-7.91E-77 3-1.03E-72 4-3.96E-72 5-4.43E-70 6-2.31E-68 7-2.71E-68 8-1.70E-66
	PD-vs.-ctrl	8202	1-Immune system process 2-Viral process 3-Cellular metabolic process 4-cell activation 5-Immune response 6-Cell activation involved in immune response 7-Myeloid leukocyte activation 8-Leukocyte activation involved in immune response	Major histocompatibility complex, class II, DQ alpha 1 (HLA-DQA1), Major histocompatibility complex, class II, DR beta 1 (HLA-DRB1), Major histocompatibility complex, class I, F (HLA-F), HLA-C, Major histocompatibility complex, class I, E (HLA-E), Exosome component 1 (EXOSC1), TIMP metalloproteinase inhibitor 1 (TIMP-1), Golgi brefeldin A resistant Guanine nucleotide exchange factor 1 (GBF1), TYROBP, SYK	1-2.52E-88 2-2.33E-87 3-1.03E-76 4-1.78E-76 5-3.10E-69 6-8.51E-62 7-2.90E-61 8-8.68E-61
	AD-vs.-PD	6599	1-Viral process 2-Cellular metabolic process 3-Cellular component organisation 4-Intracellular transport 5-Cellular protein localisation 6-mRNA metabolic process 7-Nitrogen compound metabolic process 8-Immune system process	SPEN, Voltage dependent anion channel 1 (VDAC1), C-X-C motif chemokine receptor 4 (CXCR4), Exosome component 10 (EXOSC10), Formin Like 2 (FMNL2), RAB14, SEPTIN9, Poly (RC) binding protein 2 (PCBP2), Nitrilase 1 (NIT1), AT-Rich interaction domain 5B (ARID5B), DPA1, Leucine rich repeat containing G protein-coupled receptor 4 (LGR4), Ficolin 1 (FCN1), ETS proto-oncogene 1, transcription factor (ETS1), Macrophage expressed 1 (MPEG1)	1-1.05E-83 2-7.06E-76 3-8.19E-76 4-8.41E-76 5-1.98E-74 6-8.15E-65 7-5.38E-64 8-1.93E-60

This was followed by the application of random-forest-based machine learning (ML) algorithm with the feature selection tool Boruta (Figure 2.1 D), to identify novel diagnostic and prognostic markers and therapeutic targets in both peripheral and brain tissues of NDs.

2.3.2 ML-based ranking of genes in BA9 (or DLPFC) from AD and PD

At autopsy, the moderately affected BA9 (or DLPFC) in AD and Lewy body disorders such as PD may harbour early pathogenic clues to patho-mechanisms common to both diseases. This hypothesis was tested by re-analysing RNA-seq data using the Boruta feature selection method. There were 180 features (genes) that characterised AD BA9 samples (n = 101) compared to controls (n = 69) (Figure 2.5 (A1)). The top-ranked Ontology (GO) biological processes (BPs) and

in-common genes were ‘nervous system development’ (solute carrier family 1 member 2 (*SLC1A2*), cysteine rich transmembrane BMP regulator 1 (*CRIM1*), myelin associated oligodendrocyte basic protein (*MOBP*), glycoprotein M6A (*GPM6A*), and myocyte enhancer factor 2 (*MEF2C*)), and ‘mitochondrial transport’ (heat shock protein 90 alpha family class A member 1 (*HSP90AA1*)) (Figure 2.5 (A2)). Additionally, a connection between MEF2C and GPM6A was revealed through genetic interaction and/or co-expression analysis (Figure 2.5 (A3)).

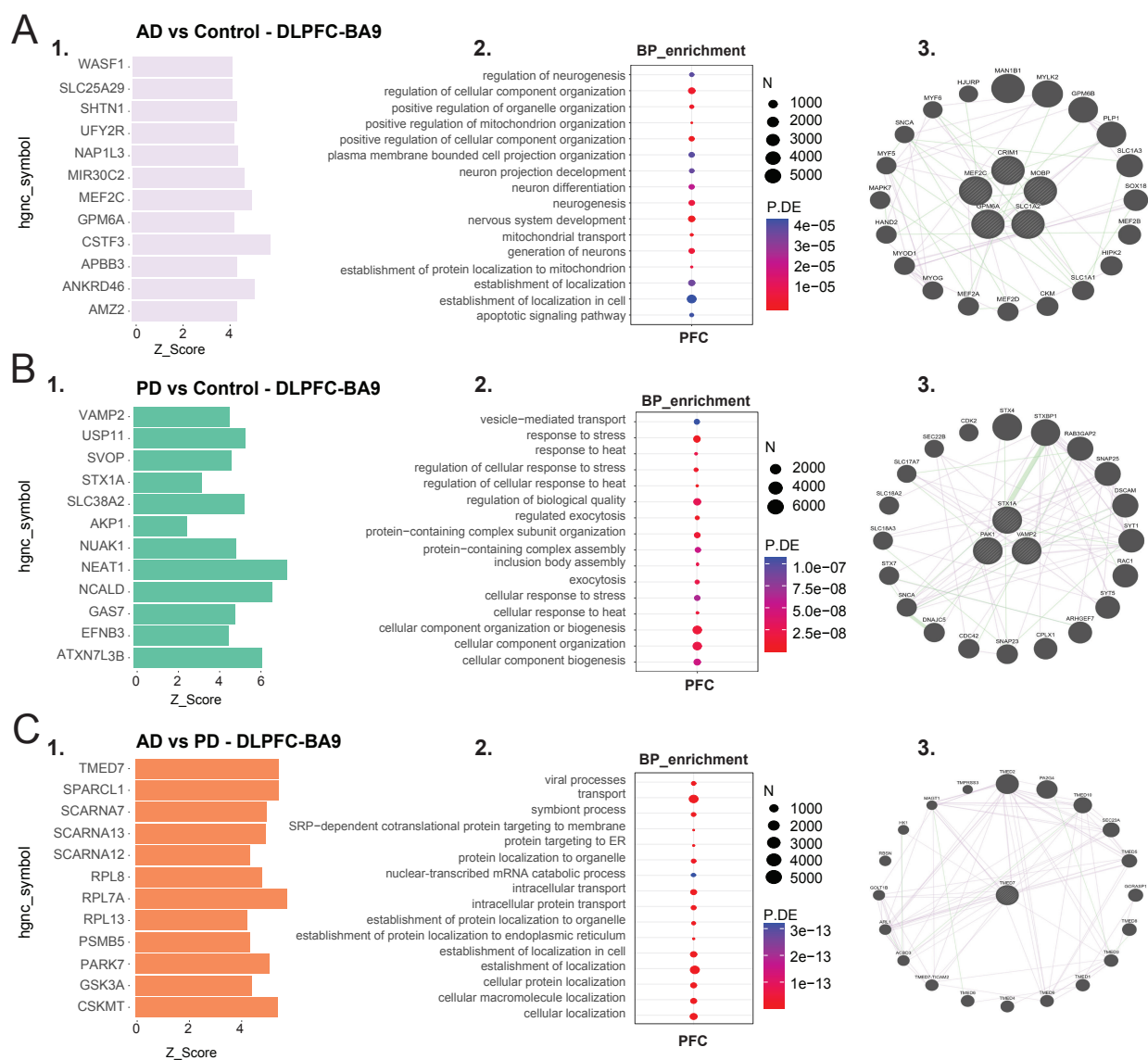


Figure 2.5 Schematic representation for DLPFC-BA9 tissue in comparing groups.

Schematic representation for DLPFC-BA9 tissue in comparing groups (A) AD-vs.-ctrl; (B) PD-vs.-ctrl, and (C) AD-vs.-PD. (1) Top twelve features (gene) selected based on their MZS shown relative to the best diagnostic marker from the feature selection tool, Boruta. (2) GO-biological process (BP) associations according to their significance for AD, PD, and NC examined in the present study. In the plots, each dot's color and size represent adjusted p -value (P.DE) and number of genes (N), respectively. (3) Gene networks that interact (green) and are co-expressed (blue) with the selected gene biomarkers (genes in the middle) from the present study.

There were 189 genes that characterised PD DLPFC/BA9 samples ($n = 126$) versus controls ($n = 93$) (Figure 2.5 (B1)). The top PD-associated GO-BPs were ‘regulation of cellular response to heat’ (BAG cochaperone 3 (*BAG3*), cysteine and histidine rich domain containing 1 (*CHORDC1*), DnaJ heat shock protein family (Hsp40) member B1 (*DNAJB1*), heat shock protein family A (Hsp70) member 8 (*HSPA8*), heat shock protein family A (*Hsp70*) member 1A (*HSPA1A*), and heat shock protein family A (Hsp70) member 1B (*HSPA1B*)), ‘exocytosis’ (P21 (*RAC1*) activated kinase 1 (*PAK1*), vesicle-associated membrane protein 2 (*VAMP2*), syntaxin 1A (*STX1A*)), and ‘regulation of cellular response to stress’ (solute carrier family 38 member 2 (*SLC38A2*)) (Figure 2.5 (B2)). A connection was shown between VAMP2, and STX1A with synaptotagmin 1 (SYT1) (Figure 2.5 (B3)). There were 235 distinct genes between AD compared to PD BA9 samples (Figure 2.5 (C1)). The top GO-BPs were ‘intracellular protein transport (transmembrane P24 trafficking protein 7 (*TMED7*))’ and ‘viral process’. Nucleophosmin 1 (NPM1), B Cell receptor associated protein 31 (*BCAP31*), and RAN binding protein 1 (*RANBP1*) were common in these two BPs (Figure 2.5 (C2)). A connection was reported between TMED7 and ADP ribosylation factor like GTPase 1 (Arl1) (Figure 2.5 (C3)).

2.3.3 Diagnosis of AD and PD by profiling peripheral blood biomarkers using ML

AD and PD-specific transcriptomic signatures were then obtained from PBMC and WB and analysed to determine if there was a feature overlap with the BA9 transcriptome for these two

NDs. Fifty-three genes characterised the AD PBMC samples (n = 22) when compared to controls (n = 14) (Figure 2.6 (A1)). Top GO-BPs were ‘cyto-skeleton-dependent intracellular transport’ (tubulin alpha 1a (*TUBA1A*), and Kinesin family member 1A (*KIF1A*)) followed by ‘viral process’ (golgi brefeldin a resistant guanine nucleotide exchange factor 1 (*GBF1*), growth factor receptor bound protein 2 (*GRB2*), ankyrin repeat domain 17 (*ANKRD17*), S-phase kinase associated protein 1 (*SKP1*), *IK*, and *FYN*) (Figure 2.6 (A2)). A connection was shown between *FYN* and zeta chain of T cell receptor associated protein kinase 70 (*ZAP70*) (Figure 2.6 (A3)). Seventy-three genes characterised PD PBMC samples (n = 6) versus controls (n = 11) (Figure 2.6 (B1)). Top GO-BPs were ‘mRNA splicing, via spliceosome’ (serine and arginine rich splicing factor 2 (*SRSF2*), heterogeneous nuclear ribonucleoprotein U-like protein 1 (*HNRNPUL1*), *IK*, and splicing factor proline and glutamine rich (*SFPQ*)) (Figure 2.6 (B2)). There was a connection between *SFPQ* and Matrin 3 (*MATR3*) (Figure 2.6 (B3)).

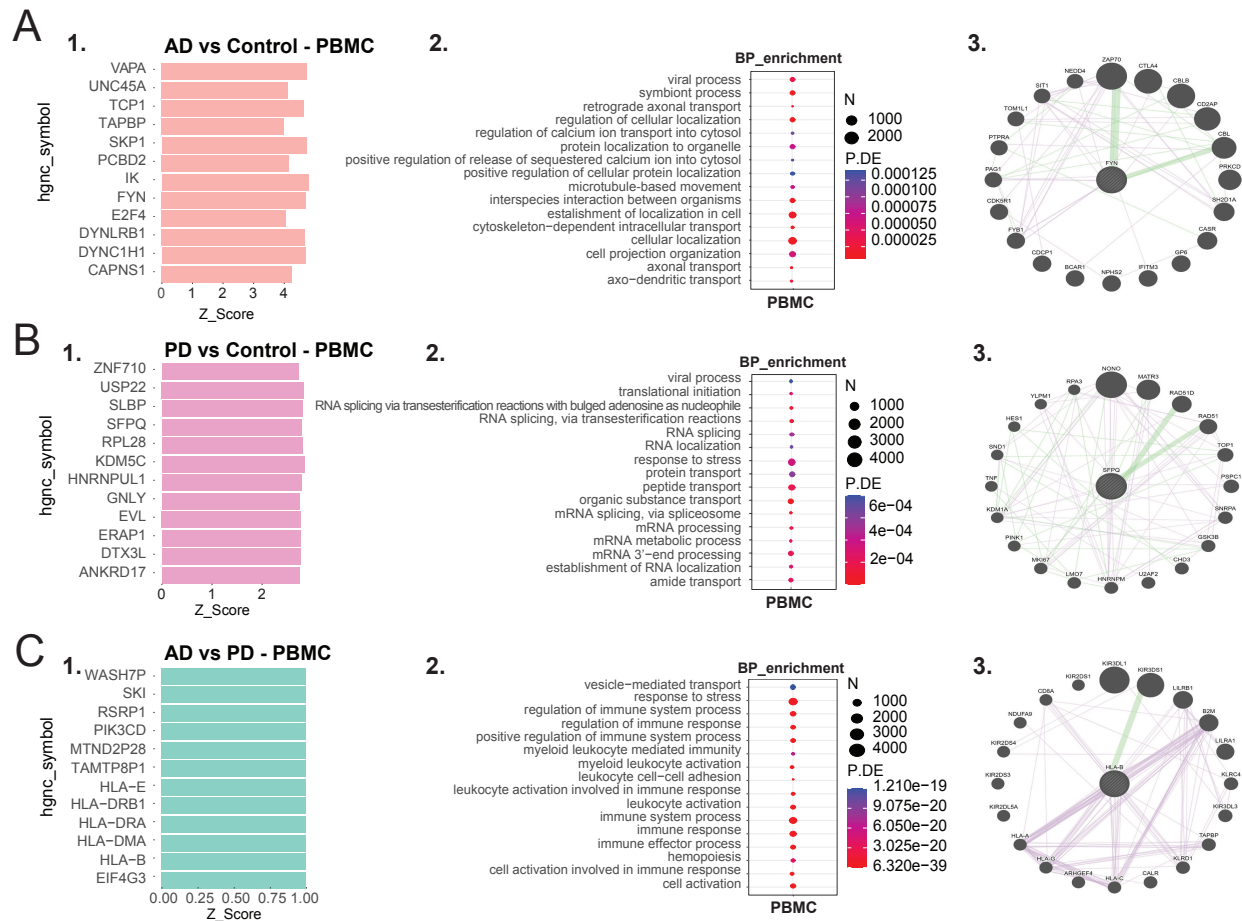


Figure 2.6 Schematic representation for PBMC tissue in comparing groups.

Schematic representation for PBMC tissue in comparing groups (A) AD-vs.-ctrl (B) PD-vs.-ctrl, and (C) AD-vs.-PD.

(1) Top twelve features (gene) selected based on their MZS shown relative to the best diagnostic marker from the feature selection tool, Boruta. (2) GO-biological process (BP) associations according to their significance for AD, PD, and NC examined in the present study. In the plots, each dot's color and size represent adjusted p -value (P.DE) and number of genes (N), respectively. (3) Gene networks that interact (green) and are co-expressed (blue) with the selected gene biomarkers (genes in the middle) from the present study.

There were 548 genes that differed between AD PBMC samples ($n = 22$) and PD ($n = 6$) (Figure 2.6 (C1)). The top GO-BP was 'immune system process' (Major histo-compatibility complex, class I, E (*HLA-E*), Major histocompatibility complex, class I, B (*HLA-B*), Major histocompatibility complex, class II, DR alpha (*HLA-DRA*), Major histo-compatibility complex, class II, DR beta 1

(*HLA-DRB1*), Major histocompatibility complex, class II, DP alpha 1 (*HLA-DPA1*)) (Figure 2.6 (C2)). Among all immune system related genes, *HLA-B* was strongly connected with killer cell immunoglobulin like receptor, three Ig domains and short cytoplasmic tail 1 (*KIR3DS1*) (Figure 2.6 (C3)).

Among WB samples, 71 genes differed between AD (n = 48) and control samples (n = 22) (Figure 2.7 (A1)). The top GO-BP was ‘interneuron migration from the subpallium to the cortex’ (ADP ribosylation factor like GTPase 13B (*ARL13B*)) (Figure 2.7 (A2)). *ARL13B* and tetratricopeptide repeat domain 26 (*TTC26*) were among the most prominent interacting genes (Figure 2.7 (A3)).

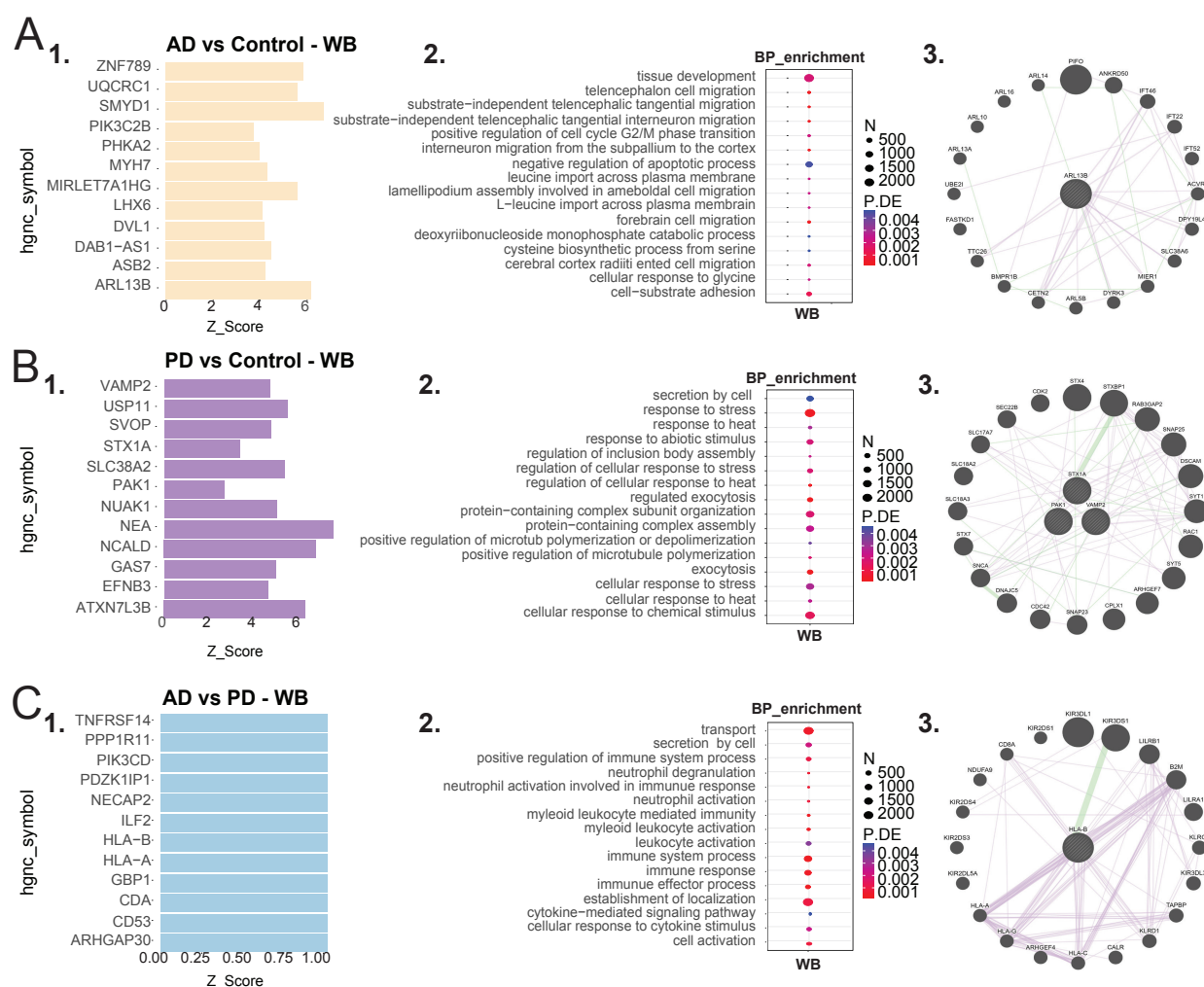


Figure 2.7 Schematic representation for WB tissue in comparing groups.

Schematic representation for WB tissue in comparing groups (A) AD-vs.-ctrl (B) PD-vs.-ctrl, and (C) AD-vs.-PD. (1) Top twelve features (gene) selected based on their MZS shown relative to the best diagnostic marker from the feature selection tool, Boruta. (2) GO-biological process (BP) associations according to their significance for AD, PD, and NC examined in the present study. In the plots, each dot's color and size represent adjusted *p*-value (P.DE) and number of genes (N), respectively. (3) Gene networks that interact (green) and are co-expressed (blue) with the selected gene biomarkers (genes in the middle) from the present study.

A total of 180 genes differed between PD WB samples (n = 20) compared to controls (n = 20) (Figure 2.7 (B1)). The top GO-BPs and genes were 'exocytosis' (*VAMP2*, *STX1A*, *PAK1*), as described above for PD-BA9 above and 'regulation of cellular response to heat' (MAPK activated protein kinase 2 (*MAPKAPK2*), *HSPA8*, *DNAJB1*, glycogen synthase kinase 3 beta (*GSK3β*), *HSPA1A*, *HSPA1B*, calcium/Calmodulin dependent protein kinase II beta (*CAMK2B*), DnaJ heat shock protein family (Hsp40) member B6 (*DNAJB6*)) (Figure 2.7 (B2)). Gene interaction and co-expression analysis were as seen in the PD-Ctrl BA9 comparison (Figure 2.7 (B3)). Lastly, there were 84 genes that differed between AD WB samples (n = 48) compared to PD (n = 20) (Figure 2.7 (C1)). The top GO-BP was 'immune system process' (*HLA-A*, *HLA-B*, and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta (*PIK3CD*)) (Figure 2.7 (C2)). This was followed by the same gene-interaction and co-expression results as seen in the PD-AD PBMC comparison (Figure 2.7 (C3)).

2.3.4 Correspondence between peripheral blood and brain in NDs

A direct comparison of blood and brain tissues for both diseases was also carried out.

There were 149 genes in common between PD-BA9 and WB samples (Figure 2.8 (A1,2)). A statistical assessment using the hypergeometric test revealed that this overlap is significant and not likely to occur by mere chance. Notably, the predominant GO-BPs for these genes were 'cellular response to heat' (*HSPA1A*, and *HSPA1B*), 'response to stress', and 'exocytosis' (*PAK1*, *VAMP2*,

and *STX1A*) (Figure 2.9 A). Gene–gene interactions were as seen in the PD-Ctrl BA9 comparison (Figure 2.9 B).

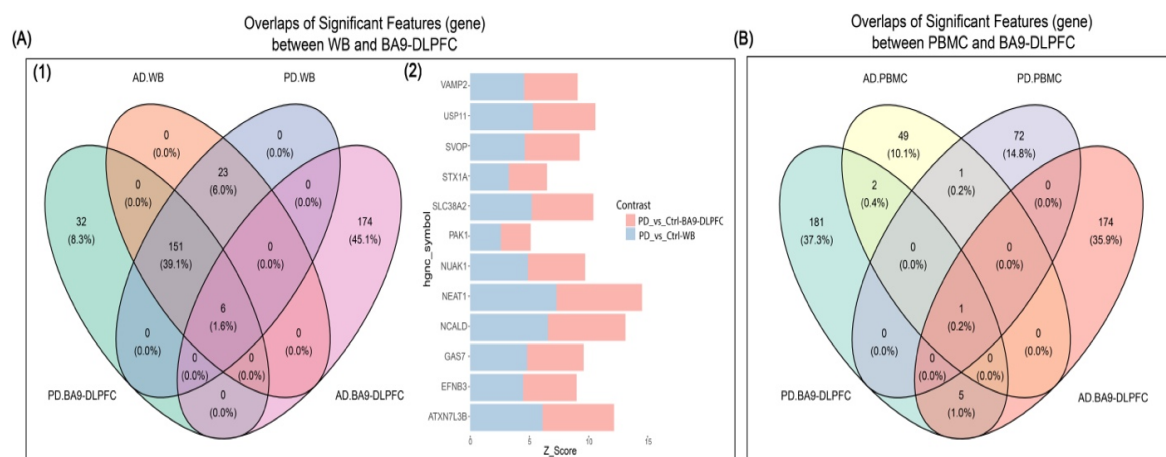


Figure 2.8 Venn diagrams illustrate the overlapping features.

(A1) Venn diagrams illustrate the overlapping features (genes) evaluated with Boruta random forest (RF)-based algorithm between Brodmann Area 9 (BA9) of the dorsolateral pre-frontal cortex (DLPFC) and whole blood (WB) samples of Parkinson's disease (PD), Alzheimer's disease (AD), and cognitively healthy controls (NC). In the diagram, colors represent the followings; WB-AD-vs.-ctrl (orange), WB-PD-vs.-ctrl (blue), DLPFC-AD-vs.-ctrl (pink), and DLPFC-PD-vs.-ctrl (green). **(A2)** The stacked plot reveals the top 12 significant features (based on their MZS) overlapped between brain (pink) and WB (blue) samples of PD patients. Vesicle-associated membrane protein 2 (VAMP2), ubiquitin specific peptidase 11 (USP11), SV2 related protein (SVOP), syntaxin 1A (STX1A), solute carrier family 38 member 2 (SLC38A2), P21 (RAC1) activated kinase 1 (PAK1), NUA family kinase 1 (NUAK1), nuclear enriched abundant transcript 1 (NEAT1), neurocalcin delta (NCALD), growth arrest specific 7 (GAS7), ephrin B3 (EFNB3), ataxin 7 like 3 (ATXN7L38). **(B)** Venn diagrams illustrate the overlapping genes between BA9-DLPFC, and PBMC samples of PD, AD, and NC. In the diagram, colors represent the followings; PBMC-AD-vs.-ctrl (yellow), PBMC-PD-vs.-ctrl (purple), DLPFC-AD-vs.-ctrl (orange), and DLPFC-AD-vs.-ctrl (green).

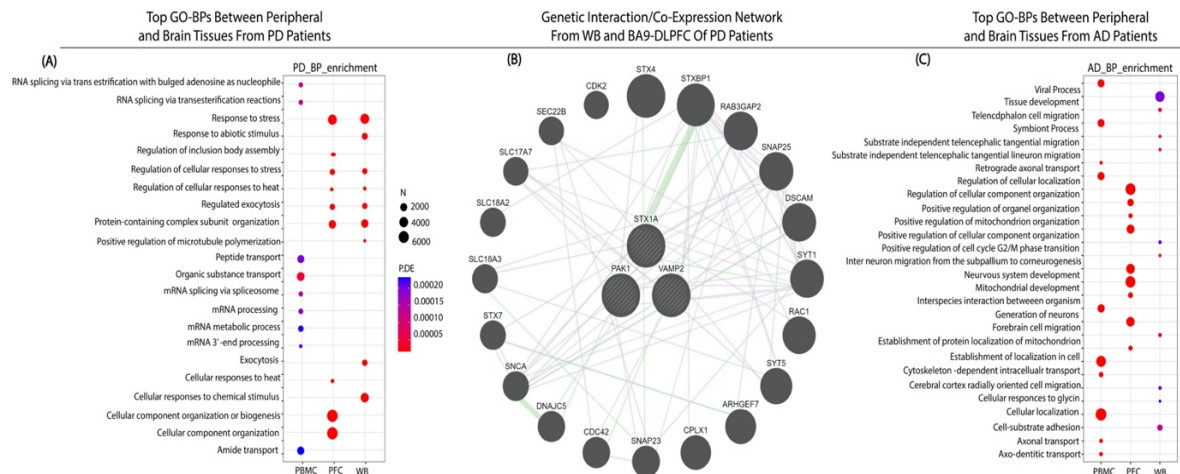


Figure 2.9 Overlapped biological processes (BP) of Gene Ontology terms.

(A) Biological processes (BP) of Gene Ontology terms that were significantly enriched for overlapped features between DLPFC/BA9 and WB samples of PD patients. In the plots, each dot's color and size represent adjusted p -value (P.DE) and number of genes (N), respectively. (B) Gene interaction (green) and co-expression (blue) network of common significant genes between central and peripheral samples of PD patients. Syntaxin 1A (STX1A), P21 (RAC1) activated kinase 1 (PAK1), vesicle-associated membrane protein 2 (VAMP2). (C) BP of gene ontology terms between blood and brain tissue of AD patients. In the plots, each dot's color and size represent adjusted p -value (P.DE) and number of genes (N), respectively.

There were no DE genes in common between DLPFC/BA9 and peripheral tissues in AD studies (Figure 2.8 B and Figure 2.9 C).

2.4 Discussions

This study explored two overlapping hypotheses that there were common transcriptomic anomalies between AD and PD, and components of the central transcriptomic signatures would also be seen peripherally. The results suggest that there is little overlap between PD and AD centrally or peripherally and that the AD brain signature is essentially unique compared to blood-derived samples. In comparison, central and 'blood' PD signatures are both characterised by anomalies in 'exocytosis' with VAMP2, STX1A, and PAK1 the most prominent genes. There is some evidence too that PD is further characterised by aberrations in immune processes.

VAMP2 and STX1A, critical proteins in the neuronal SNARE complex, were the prominent genes in the PD-BA9. VAMP2, a vesicle-associated (v)-SNARE, plays a crucial role in Ca^{+2} -dependent exocytosis of synaptic vesicles, while STX1A, associated with the plasma membrane ((t)-SNARE), is specifically involved in vesicle fusion [172]. Interaction between v-SNARE (VAMP2) and t-SNARE (STX1A) induces membrane bending forces, which results in the pulling together of synaptic vesicles (SVs) and presynaptic membranes, thereby facilitating membrane fusion and exocytosis events [173]. Any dysregulation in their expression can reduce active fusion vesicle counts, aligning with the role of α -synuclein (α -syn) in modulating SNARE assembly [173, 174]. Pathologically, studies in PD mice have revealed an interaction of large α -syn oligomers with the N-terminus of multiple VAMP2s on vesicles. This incapacitates v-SNARE's (VAMP2) ability to interact with t-SNARE (STX1A), inhibiting SNARE complex formation at the synaptic site and thus blocking exocytosis [152]. Interestingly, VAMP2 and STX1A were co-expressed with SYT1. SYT1 is a major Ca^{2+} -sensor protein that regulates synaptic vesicle exocytosis through direct interactions with VAMP2 and STX1A after Ca^{2+} influx in the presynaptic terminal [175]. Reduced levels of synaptic vesicle proteins (SYT1, VAMP2, and STX1A) along with increased expression of α -syn have been associated with synaptic dysfunction and neuronal degeneration in PD mice models [176].

A second finding here is that human leukocyte antigens, specifically the Class I molecules of the B gene (HLA-B), differentiate AD from PD blood-based transcriptomes. HLA-B significantly interacts with KIR3DS1. KIR3DS, a member of kill-er-immunoglobulin-like receptors (KIRs) is present on the surface of natural killer (NK) cells and can bind to HLA-B present on the surface of all nucleated cells. These interactions modulate NK cell activity, including the killing of virus-infected cells and tumours, or induction of cytokine secretion [177]. Changes in HLA-B7 activity,

along with a decrease in cytotoxic function of NK cells have been implicated as significant contributors to late-onset AD (LOAD) in subjects without apolipoprotein E (APOE) $\epsilon 4$ [178-179]. The most prominent feature of the AD- versus PD-BA9 comparison was TMED7, a type I transmembrane protein of the p24 protein family of the early secretory pathway, which is involved in the regulation of innate immune signalling. In a previous study, TMED7 was shown to facilitate myeloid differentiation marker 88 (MyD88)-dependent toll-like receptor 4 (TLR4) signalling by forming a stable complex with its ectodomain and thus promoting TLR4 translocation from the Golgi to the cell surface through its vesicular trafficking activity [180]. Additionally, TMED7 inhibits MyD88-independent TLR4 signalling by disrupting the TRIF-related adaptor molecule (TRAM)/TIR domain-containing adaptor protein inducing interferon β (TRIF) complex in late endosomes [181]. Although an association with NDs has not yet been established for TMED7, reduced levels of trans-membrane protein 21KD (TMP21)/TMED10, a member of the p24 cargo receptor family, are known to be associated with AD progression [182, 183]. Gene network analysis showed an interaction between TMED7 and ADP-ribosylation factor-like protein 1 (Arl1), a member of the ARF/Arl family of small GTPases localised on trans-Golgi membranes [184]. Previous studies showed that depletion of Arl1 causes dissociation of golgin-97, and golgin-245 is anchored to the trans-Golgi membrane via a carboxy-terminal GRIP domain, which binds to membrane-associated Arl1 [184, 185]. This impairs both vesicle trafficking and the Golgi integrity at the trans-Golgi network (TGN), which causes Golgi fragmentation, alteration of Golgi positioning, and impaired secretory traffic [184, 186]. This is interesting since both TMED7 and Arl1 are involved in Golgi vesicle-mediated transport, and its alteration is likely to induce Golgi fragmentation in neurodegenerative diseases, including AD and PD [187].

Mef2C, a transcription factor implicated in the regulation of innate and adaptive immune cells with a myeloid origin, including microglia, was prominent in the AD-NC DLPFC/BA9 comparison. In the CNS, Mef2C has been shown to restrict microglial responses to immune stimuli by functioning as an ‘off’ switch [188]. Studies on aged mice models have suggested that Mef2C loss of function mediates chronic elevation of type I interferon (IFN-I) response related to early microglial activation in AD-related conditions [188, 189]. Mef2C is also involved in neuronal formation and differentiation, as well as in the growth and pruning of axons and dendrites through interaction with GPM6A. GPM6A is a member of the tetraspan proteolipid proteins (PLP) and a direct target gene for Mef2C that has been shown to be co-expressed with Mef2C [190].

Although not directly overlapping here, peripheral PBMC signatures did reveal genes that had previously been shown to be associated with AD in brain tissue analyses. Increased levels of FYN in AD brains have been proposed to regulate amyloid precursor protein (APP), phosphorylation at tyrosine 682 (Tyr682) residue, and be involved in oligomeric amyloid- β (A β)-mediated synaptic toxicity [191]. In turn, oligomeric A β increases local translation of the axonally-enriched protein Tau in the somatodendritic domain via activation of FYN/MAP kinase (MAPK)/ribosomal protein S6 signalling pathway, thereby linking the two molecules that accumulate in AD brains [192]. Additionally, genetic interaction/co-expression network analysis revealed a connection between FYN and zeta-chain (TCR)-associated protein kinase, 70-kDa (ZAP70). Up-regulation of these two kinases is thought to increase peripheral mast cell and B-/T-cell activation, thus suggesting a role for peripheral blood immune dysregulation in AD [193, 194].

A gene that specifically characterised the peripheral PD transcriptome encodes SFPQ, also known as polypyrimidine tract binding protein-associated-splicing factor (PSF), with roles in DNA transcription and repair and RNA processing. Previous PD animal model studies have reported

dysregulated interactions between SFPQ/PSF and LIM homeobox transcription factor 1 beta (Lmx1b)/nuclear receptor-related 1 protein (Nurr1), which results in metabolic impairment, α -syn inclusions, and progressive loss of dopaminergic neurons [195, 196]. Additionally, disrupted interactions between SFPQ/PSF and fused in sarcoma (FUS) has been shown in postmortem frontotemporal lobar de-generation (FTLD) brains and were associated with increased 4-repeat tau (4R-T) to 3-repeat tau (3R-T) ratios and underlying FTLD phenotype development. This seems likely due to dysregulated alternative splicing of microtubule-associated protein tau (MAPT) gene at exon 10 [197]. Additionally, our studies showed an interaction between SFPQ and MATR3, which is involved in DNA damage response (DDR) and activated by DNA double-strand breaks (DSB). MATR3 and SFPQ depletion, along with increased levels of phospho- α -syn are associated with the accumulation of DNA DSBs that contribute to programmed cell death in in vitro models of PD [198, 199].

One gene that defined AD-WB samples was ARL13B, a member of the Ras family of GTPases with distinct regulatory roles in primary cilia protein trafficking, the Sonic Hedgehog (Shh) pathway, and neural development [200]. Studies on AD mouse models have revealed mislocalisation of ARL13B on primary cilia membranes, a microtubule-based sensory organelle present in neurons and astrocytes. This led to defective cilia-mediated TLR4/NF- κ B activation, reduced axonal length, and altered signal transmission [201, 202]. Our analysis showed no genes interacting with ARL13B. However, ARL13B is co-expressed with TTC26. TTC26, also known as intraflagellar transport 56 (IFT56), which is responsible for ARL13B localisation on the cilia membranes. Reduced levels of TTC26 lead to ARL13B mislocalisation and impaired Shh signalling pathway associated with abnormal maintenance of ciliary structure [203, 204].

Overall, there is little overlap in either the central or peripheral transcriptomes of the two most common NDs, AD and PD. Unlike AD, there is an overlap between peripheral and central signatures of PD, suggesting anomalies in exocytosis and immune function. Thus, PD blood biomarkers may have more potential for monitoring the severity and progression of the disease. In the absence of blood biomarkers, the AD-BA9 signature reinforces the idea of anomalies in innate immune signalling and specifically TLR4/NF- κ B activation.

2.5 Conclusions

Applying a machine learning algorithm to existing central and peripheral transcriptomic has clarified that there is little overlap between the two diseases. The PD-DLPFC/BA9 was characterised by alterations of SNARE proteins (VAMP2 and STX1A) which negatively affect neurotransmitter release at the axon terminal, while AD-BA9 was characterised by changes in ‘nervous system development’ and MEF2C, a transcription factor, potentially inducing microglia activation. Analyses of a gene signature from PBMC samples of AD subjects were associated with alterations in ‘viral process’ and FYN, which is associated with A β and tau in the brain but without overlap with the central transcriptome. It is unclear whether peripheral levels of FYN reflect disease severity and progression.

Whereas the central and peripheral PD-transcriptomes were both characterised by anomalies in exocytosis and particularly the differential expression of genes associated with the SNARE complex. This is consistent with our current understanding of the physiological role of α -syn and how oligomers may compromise vesicle docking and neurotransmission.

2.6 Potential limitations and future prospective

Results from the present study have several notable limitations. Firstly, I could have missed relevant disease-associated gene sets across studies due to specific study designs, within-study

biases, and variation across studies. However, as the scientific landscape evolves, so do the tools at our disposal. For instance, technological advancements like long nanopore reads are transforming the field of gene sequencing. By enabling the sequencing of full-length transcripts from start to finish, nanopore techniques ensure their precise and clear-cut analysis. This provides a deeper understanding of gene expression and sheds light on its diverse isoforms.

Secondly, although I selected nearly 330 independent human peripheral RNA-seq studies for inclusion in our analyses, prioritising those most relevant to AD, PD, and cognitively healthy controls, I omitted many others with the potential to provide additional insights. In the future, our approach can thus be generalised to an even broader sample of available human peripheral data. Thirdly, through our analysis, a number of potentially essential genes were detected that remained incompletely defined and were likely to influence AD and PD. Therefore, it will be important in future work to consider gene sets significantly enriched for disease-specific SNPs through genome-wide association studies (GWASs). Fourthly, human RNA-seq data derived from bulk brain tissue include mixed cell type composition, which may therefore reflect global changes in immunity and neuronal pathways. The growing availability of single-cell expression profiles can definitively address this concern in future work.

Additionally, as users of public data in which all cases are sporadic and do not include mutation carriers' information, I expect heterogeneity, i.e., different patients having their combination of genetic and lifestyle factors. Still, this study is limited to finding common factors based on the assumption that all cases have a core of causative factors or a body of brain (expression) changes in common. Finally, the relatively shallow sampling in some comparisons and the prevalence of immunological gene ontology terms raise questions about the influence of patients' immune status on the observed enrichment. For instance, PD patients may be more prone to physical injuries and,

consequently, more infections, while control patients might be more mobile and exposed to greater microbiological diversity than home-bound ND patients. These factors could potentially confound the results, making it challenging to confidently attribute the observed enrichment solely to the diseases under study. Future works should encompass genotype data, more thorough sampling, and the integration of proteome sequencing to enhance our grasp on the genetic, environmental, and post-transcriptional facets influencing these conditions.

Chapter 3 Transcriptome based genetic variant discovery from peripheral and central tissues to characterise Alzheimer's and Parkinson's disease

3.1 Introduction

Clinically, Alzheimer's disease (AD) and Parkinson's disease (PD) occur in familial and sporadic forms and are characterized by progressive motor and non-motor symptoms. The identification of various genetic mutations responsible for familial forms of NDs has enhanced our understanding of the molecular processes underlying disease pathogenesis. For instance, mutations in genes such as apolipoprotein E (APOE), presenilin 1 (PSEN1), and PSEN2 are implicated in AD, while synuclein alpha (SNCA), PTEN-induced kinase 1 (PINK1), parkin (PRKN), leucine-rich repeat kinase 2 (LRRK2), and protein DJ-1 (DJ-1) have been associated with PD. However, most of these cases occur as sporadic forms, likely resulting from uncharacterized genetic variants and environmental factors interplay. Despite many years of research and remarkable progress in the knowledge of the molecular pathogenesis of NDs, a full spectrum of the sporadic form disease-associated variations is still not yet within reach. This would limit the understanding of AD and PD pathology and study of future therapeutic targets.

Recently, RNA sequencing (RNA-seq) has become a powerful tool for the study of differential gene expression patterns and their potential functional impacts on diseases or response to environmental changes. Additionally, there are few reported studies which have used RNA-seq reads with no matched WGS or WES data for the identification of genetic variants required for analyzing the association of polymorphisms with traits [205-207]. Through correct mapping of RNA-seq reads to the GRCh38 reference genome (GRCh38.p12.genome.fa) and application of computational pipelines for variant calling and vigorous filtering of false-positive calls; RNA-seq data can therefore provide sets of numerous and well distributed single nucleotide polymorphisms (SNPs) throughout the genome [208].

This study aimed to set up a procedure allowing SNP detection from sequence read archive (SRA)-RNA-seq data utilizing a production bioinformatics pipeline (https://github.com/binfnstats/RNA-seq_SNP-calling). Then, detected SNPs were used for case-control and case-case genome-wide association study (GWAS) to reveal variants differentially associated with AD, and PD. This was followed by an analysis strategy based on SNPs-genes-enriched pathways to identify genetic mechanisms important for characterizing the most common NDs.

3.2 Materials and methods

3.2.1 Data collection

Publicly available RNA sequencing (RNA-seq) studies of (i) human post-mortem brain tissue and (ii) peripheral blood from patients with AD, PD, and cognitively healthy controls (NCs) were downloaded from the sequence read archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) database using the getFASTQfile function from the SRADB R/Bioconductor package [161].

RNA-seq data were available for 763 samples from postmortem brain region: Brodmann area 9 only (BA9) (n = 133 with AD, n = 171 with PD, and n = 168 NC), and blood: PBMCs datasets (n = 22 with AD, n = 46 with PD, and n = 55 NC) and WB (n = 53 with AD, n = 38 with PD, and n = 77 NC).

For the comparison of SNPs detected by RNA-seq versus DNA-seq, genomic-based SNPs associated with AD were accessed from “Bellenguez et al. 2022. Nature Genetics” with n= 2,447 AD diagnosed cases, n= 46,828 proxy cases of dementia and n= 338,440 controls.

Additionally, genomic-wide SNPs associated with PD were downloaded from “Nalls et al. 2019. Lancet Neurology”, which includes the data obtained from three previously published studies: “Nalls et al. 2014” with n= 13,708 PD diagnosed cases and n= 95,282 control, “Chang et al. 2017” with n= 6,476 PD cases and n= 302,042 controls, and the publicly available “neuroX” dataset from the

international Parkinson's disease genomics consortium (IPDGC), whose dataset consisted of n= 6,584 PD cases and n= 15,470 controls.

3.2.2 RNA-seq mapping and data processing

The alignment of raw reads into a reference genome was the first and essential step in detecting genetic variants. *Spliced transcripts alignment to a reference* (STAR) version 2.7.9a [209] with “—twopassMode” option was the alignment tool used for mapping RNA-seq reads against the GRCh38 reference genome (GRCh38.p12.genome.fa). The pipeline (https://github.com/binfstats/RNA-seq_SNP-calling/blob/main/Scripts/paired-star-align.wdl) for STAR alignment was written in the workflow description language (WDL) that utilizes Docker images, and runs in Cromwell server.

3.2.3 RNA-based variant detection using GATK pipeline

The genome analysis toolkit (GATK) Best Practices pipeline (<http://www.broadinstitute.org/gatk>) was utilised for variant discovery and genotyping from RNA-seq data. The GATK RNA pipeline comprises a series of optional steps that have been demonstrated to enhance the quality of variants derived from high throughput sequencing [210]. These steps can be categorized into pre-processing, variant calling, and variant filtering stages, with each stage having specific functions. The pre-processing stage involves steps to prepare the RNA-seq data for accurate variant calling. The scripts for this stage are available at https://github.com/binfstats/RNA-seq_SNP-calling/blob/main/Scripts/Pre-processing-picard.sh. In the study, the following Picard tools (<http://broadinstitute.github.io/picard/>) were employed for pre-processing:

“--ReorderSam”: To reorder reads in a SAM/BAM file to match the contig ordering in the provided reference genome. “--AddOrReplaceReadGroups”: To insert read group information into the reads. “--MarkDuplicates”: To identify and mark duplicate reads, mitigating potential biases in variant calling.

Variant calling encompasses steps to identify variants in the aligned reads. The scripts for this stage can be accessed at https://github.com/binfnstats/RNA-seq_SNP-calling/blob/main/Scripts/Variant-calling-GATK.sh. The following GATK functions was employed for variant calling:

“--SplitNCigarReads”: To split reads that overlap introns. “--BaseRecalibrator” and “--ApplyBQSR”: To recalibrate bases with mapping quality scores. “--SplitIntervals”: To create specific genomic intervals (subsets of genomic regions) for parallelization. “--HaplotypeCaller”: To perform variant calling for each sample using the option “--stand_call_conf 20”.

The variant filtering stage involves steps to filter the called variants based on specific criteria. The scripts for this stage can be found at https://github.com/binfnstats/RNA-seq_SNP-calling/blob/main/Scripts/VCFs-merge-filter.sh. The following steps for variant filtering were utilised:

Merge: Variants from multiple samples were combined into a single variant calling format (VCF) file using bcftools (v.1.9) with the “--merge” option. Variant filtration: Variants were filtered using the GATK function “--VariantFiltration” with the filters “--QD < 2” and “--FS > 30”.

3.2.4 Annotation and distribution of variants across the genome

To analyse, annotate, and partition detected SNPs in different genomic regions, the ensembl variant effect predictor (VEP) toolset [211] with default parameters was used. According to functional annotations of total SNPs, they were partitioned into 16 different categories; including intron variant, splice region variant, splice donor variant, splice acceptor variant, non-coding transcript exon variant, non-coding transcript variant, MND-transcript variant, 5′ and 3′ untranslated regions (UTRs) variant, upstream gene and downstream gene variant, intergenic variant, missense variant, synonymous variant, regulatory region variant, start site retained variant and stop site retained variant. The current study excluded all splice variants and start and stop sites because they contained a very low proportion of the total SNPs.

3.2.5 Quality control (QC) and genome-wide association studies (GWAS)

QC and GWAS procedures were implemented in the free, open-source whole-genome association analysis toolset PLINK version 1.9 [212] with default parameters and R statistical software. Following sample quality control, individuals with a heterozygosity rate $\leq 0.03\%$, and sex discrepancy were excluded from our study. SNPs with poor genotype clustering performance, excess missingness $> 1\%$, minor allele frequency (MAF) $\leq 1\%$, or out of Hardy Weinberg equilibrium (HWE; $p \leq 1e-6$) were removed. Duplicated variants were also excluded from our study. This was followed by a standard association analysis between variants and complex traits using “--assoc” command with “--condition” and “--pfilter 1e-1” flag to provide either asymptotic (likelihood ratio test and Wald test) or empirical significance P -values.

SNP density, Quantile-quantile (Q-Q) and Manhattan plots of the marginal asymptotic P -values were constructed using the R Bioconductor package “CMplot”.

Next, the PLINK clumping procedure was used to define significant ‘hits’ in each region. An iterative clumping procedure was applied to all variants with a p -value below 1×10^{-5} , starting with the variant with the lowest P -value (referred to as the index variant). Variants with a p -value below 1×10^{-5} , located within 250 kb of this index variant and in LD with the index variant (r^2 above 0.1) were assigned to the index variant’s clump. The clumping procedure was then applied until all the variants had been clumped.

3.2.6 Locus gene-based pathway analysis

Pathway scoring algorithm (PASCAL) [213], an open-source software package, was used to provide gene scoring by aggregating SNP p -values from GWAS association results while correcting for linkage disequilibrium (LD) across genomic regions using 1000 Genomes Project data.

To extract biological insights from genes of interest, a WEB-based gene set analysis toolkit (WebGestaltR) [214], an R Bioconductor package, was used. WebGestaltR enables more comprehensive, powerful, flexible, and interactive functional enrichment analysis for interpreting genome-wide expression profiles in comparison with other methods.

3.3 Results

3.3.1 Genomic variant detection from RNA-seq data workflow

The genomic variant detection workflow began with the acquisition of RNA-seq data from the Sequence Read Archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>), as shown in Figure 3.1 A and Appendix 3. Following data acquisition, the RNA-seq reads were aligned to the GRCh38 reference genome using the STAR aligner with the “--twopassMode” method, as depicted in Figure 3.1 B. After alignment, a series of pre-processing measures were carried out using the Picard tools (<http://broadinstitute.github.io/picard/>), as detailed in Figure 3.1 C. Upon completing the pre-processing steps, the GATK RNA pipeline was applied for variant calling and filtering as shown in Figure 3.1 C.

Next, stringent quality control (QC) was applied (Figure 3.1 D) to exclude SNPs that had low genotype call rates (< 0.95) or did not follow HWE. Subjects with low sample call rates (< 0.95), incompatible sex, cryptic relatedness, or duplication across cohorts were also excluded. This was followed by the application of a genome-wide association study (GWAS) (Figure 3.1 D) to identify associations between the genotype at each SNP and case-control/case-case status using either asymptotic (likelihood ratio test and Wald test) or empirical significance values.

In addition, association tests at the levels of genes and pathway were conducted (Figure 3.1 E). The association results would be strengthened if disease correlates of these variants could be found.

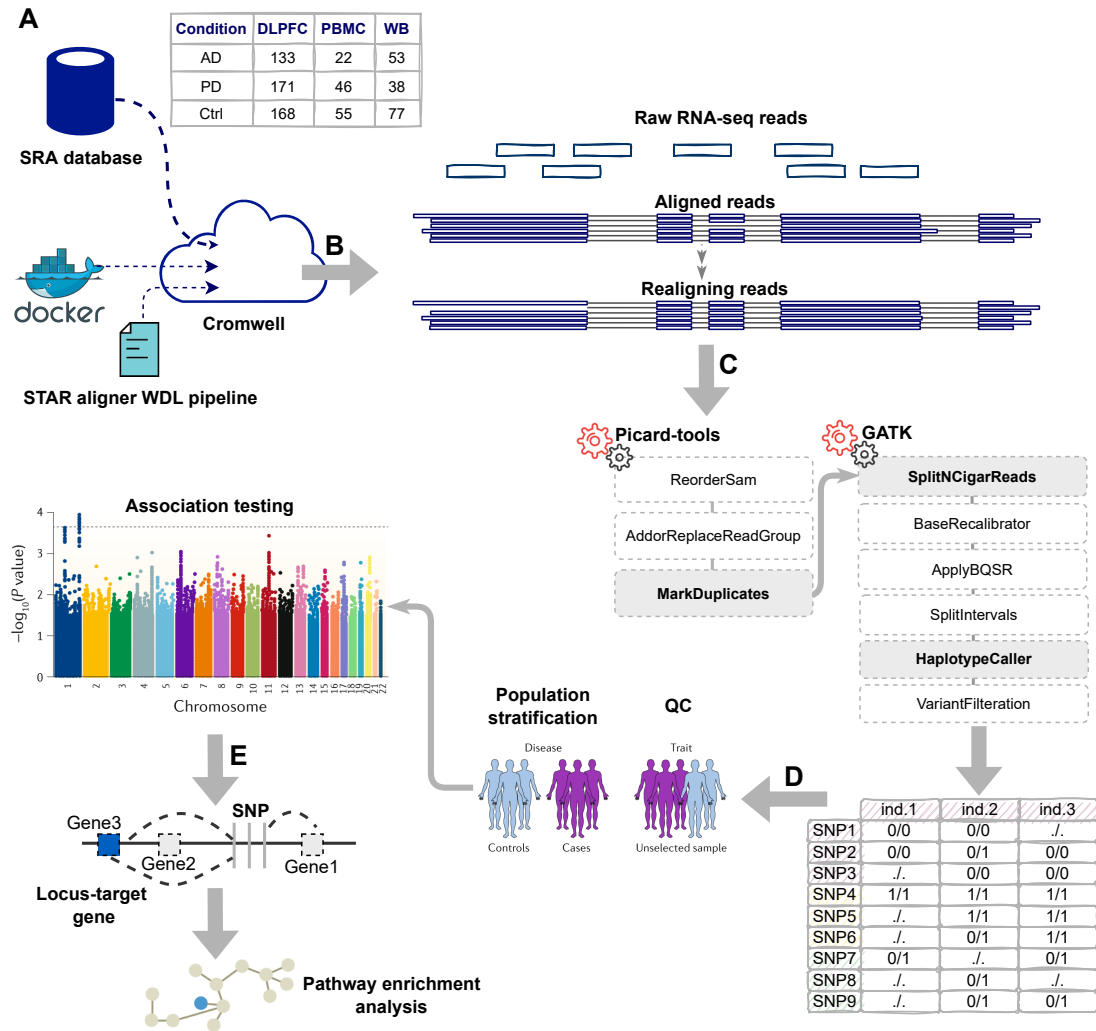


Figure 3.1 Schematic representation showing all steps necessary for RNA-seq based genomic variant discovery and downstream analysis.

(A) Workflow for the discovery of genomic variants by RNA-seq reads begins with a search in SRA database for RNA-seq data to access publicly available human post-mortem brain- and blood-based studies. Next, raw RNA-seq reads were used as the input for the alignment pipeline, performed using a “--twopassMode” method with STAR. The pipeline was wrapped into workflow description language (WDL) and was executed in cloud-based services with Cromwell and Docker combination. (B) Through the alignment workflow, STAR aligner performed the 1st pass mapping. Then it automatically extracted junctions, inserted them into the genome index, and, finally, re-mapped all reads in the 2nd mapping pass. This allows to detection of more splice junctions in reads mapping to novel junctions. (C) Uniquely mapped reads were pre-processed (using Picard tools) following the GATK best practices for detecting genomic variants from RNA-seq data. (D) Filtered variants were passed quality-control (QC) and tested for association

using PLINK. (E) Finally, SNPs with the nominal significance of association were mapped to genes if they were located within 50 kb upstream or downstream of genes' transcription start sites' using PASCAL. A gene-set enrichment analysis followed this to identify associated biological pathways.

3.3.2 Replication studies for SNPs detected by RNA-seq with those detected by genome sequencing (not same individuals)

In the current research, with the filters suggested by GATK pipeline, the number of SNPs detected with RNA-seq reads were as follows: AD samples (DLPFC/BA9 (2,953,832), PBMC (2,940,972), WB (3,831)) and PD samples (DLPFC/BA9 (2,661,516), PBMC (508,959), WB (296,308)).

To assess replication of SNPs detected by RNA-seq data, genomic-based SNPs from independent studies were accessed for AD (30,297,401) and PD (17,510,618) (see materials and methods for details). In the PD study, genomic SNPs were built among human genome reference 19 (hg19). To convert single nucleotide variants between genome builds (hg19 to hg38), SNPnexus (<https://www.snp-nexus.org/v4/>), a web-based annotation tool, was used. After conversion 8,310,764 variants were left.

In the current RNA-based SNP detection study, the number of replications with SNPs reported from genomic studies were as follows: AD samples (DLPFC/BA9 (2,661,516), PBMC (2,629,458), WB (3,515)) and PD samples (DLPFC/BA9 (42,840), PBMC (11,853), WB (8,086)). Assuming SNPs detected by genome sequencing represent the “truth”, these numbers represent the sensitivity, or recall, of RNA-seq for SNP detection.

As expected, the number of SNPs detected with RNA-seq is much lower than that in genome sequencing data, because only variants present in transcribed regions were detected. Importantly, the small sample size used in the current study “specially for peripheral blood” might affect both coverage and power to detect more disease associated variants. Another reason could be that there are differences between the populations used in the genomic studies and the population used in the

current study. Finally, lower number of replications reported in PD studies in comparison to AD might be related to the conversion stage of genomic SNPs, where many possible SNPs were lost.

3.3.3 Genomic location of SNPs detected by RNA-seq for each comparison

Next, the repartition of the SNPs detected by RNA-seq across different genomic regions (Figure 3.2) was studied for each case-control and case-case analysis.

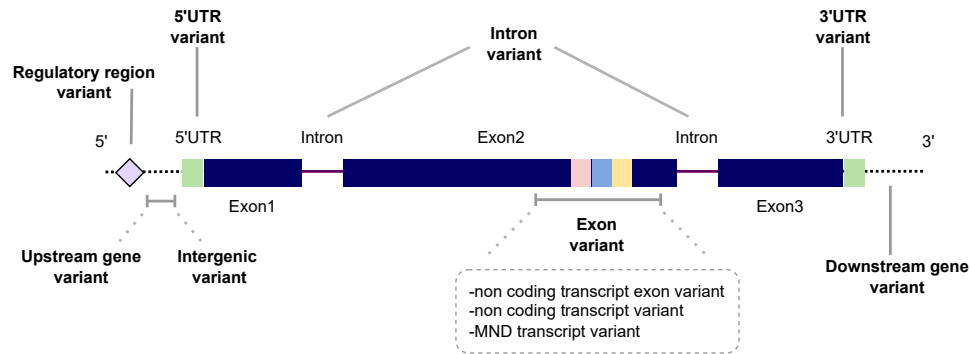


Figure 3.2 Repartition of the SNPs detected by RNA-seq among different genomic regions.

Figure illustrates the genome partitioning of genetic variants (SNPs) found by RNA-seq data. Detected SNPs were mainly distributed in intronic, and exonic regions (5' and 3' untranslated regions (UTRs), non-coding transcript exon, non-coding transcript, and MND transcript region). This was followed by SNPs located at 5-kb upstream and 5-kb downstream of genes. lower proportion of the variants were in the intergenic (SNPs located in the regions between two adjacent genes in the genome) and regulatory regions.

There were 3,731,887 variants called from AD DLPFC/BA9 samples (n= 133) merged with controls (n= 60) (Figure 3.3). After QC only 2,001,311 variants were left. The majority of the QC variants were located in the intronic region (56%) followed by exonic regions (28%), downstream gene (6%), upstream gene (5%) and at lower proportions in regulatory (2%) and intergenic (1%) regions.

There were 2,205,121 variants from PD DLPFC/BA9 samples (n= 171) merged with controls (n= 108) (Figure 3.3). After QC only 909,247 variants were left, from which most of the variants were

in the intronic region (52%) followed by exonic regions (27%), downstream gene (9%), upstream gene (7%) and at lower proportions in regulatory (2%) and intergenic (1%) regions.

There were 3,410,466 variants called from AD DLPFC/BA9 samples (n= 133) merged with PD (n= 171) (Figure 3.3). After QC 1,748,877 variants were left, from which many of the variants were located in the intronic region (55%) followed by exonic regions (27%), downstream gene (7%), upstream gene (6%) and at lower proportions in regulatory (2%) and intergenic (1%) regions.

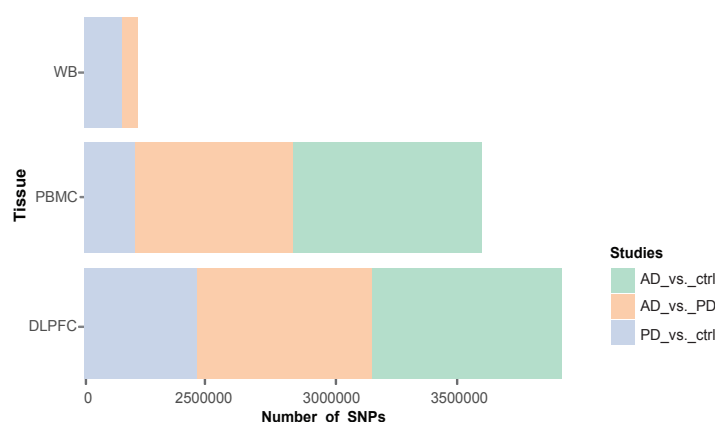


Figure 3.3 Total number of SNPs detected by RNA-seq among comparing groups.

The number of SNPs discovered by RNA-seq are shown for AD_vs_ctrl (green), PD_vs_ctrl (purple), and AD_vs_PD (orange) studies in DLPFC/BA9, PBMC, and WB samples.

There were 3,682,870 variants called from AD PBMC samples (n= 22) merged with controls (n= 14) (Figure 3.3). After QC only 1,825,881 variants were left from which most of the variants were in the intronic region (56%) followed by exonic regions (27%), downstream gene (7%), upstream gene (6%) and at lower proportions in regulatory (3%) and intergenic (1%) regions.

There were 981,578 variants called from PD PBMC samples (n= 46) merged with controls (n= 41) (Figure 3.3). After QC only 386,233 variants were left, from which many of the variants were located in the intronic region (51%) followed by exonic regions (25%, downstream gene (10%), upstream gene (8%) and at lower proportions in regulatory (2%) and missense (1%) regions.

There were 3,103,082 variants called from AD PBMC samples (n= 22) merged with PD (n= 46) (Figure 3.3). After QC only 1,401,604 variants were left, from which the majority of the variants were in the intronic region (55%) followed by exonic regions (27%), downstream gene (7%), upstream gene (6%) and at lower proportions in regulatory (3%) and intergenic (1%) regions.

There were 6,852 variants called from AD WB samples (n= 53) merged with controls (n= 32) (Figure 3.3). After QC only 2,944 variants were left, from which the majority of the variants were located in the intronic region (30%) followed by exonic regions (27%), downstream gene (12%), upstream gene (11%) and at lower proportions in missense (9%) and synonymous (6%) regions.

There were 738,320 variants called from PD WB samples (n= 38) merged with controls (n= 45) (Figure 3.3). After QC only 728,366 variants were left, from which most of the variants were in the intronic region (48%) followed by exonic regions (24%), downstream gene (12%), upstream gene (9%) and at lower proportions in regulatory (2%) and missense (1%) regions.

There were 298,412 variants called from AD WB samples (n= 53) merged with PD (n= 38) (Figure 3.3). After QC only 144,208 variants were left, from which many of the variants were located in the intronic region (39%) followed by exonic regions (25%), downstream gene (18%), upstream gene (11%) and at lower proportions in regulatory (2%) and missense (2%) regions.

These findings indicate that the SNPs identified via RNA-sequencing were primarily distributed within intronic regions, followed by exonic regions. These areas play critical roles in gene expression, alternative splicing, and chromatin structure. Notably, variations in intron- and exon-specific SNPs have been linked to the risk of developing AD and PD. It is also important to highlight that while SNPs detected through RNA-seq predominantly represent transcribed regions, those identified in intergenic spaces may reside within unannotated genes or regions of genes not yet catalogued, especially within the 3'UTR and 5'UTR regions.

3.3.4 Genome-wide association studies (GWAS) in BA9 (or DLPFC) from AD and PD

Next, association analysis was performed on QC variants to analyse the genetic correlation in each comparison. Under the null distribution QQ plots of calculated p -values for the discovery phase of all traits were calculated (Figure 3.4).

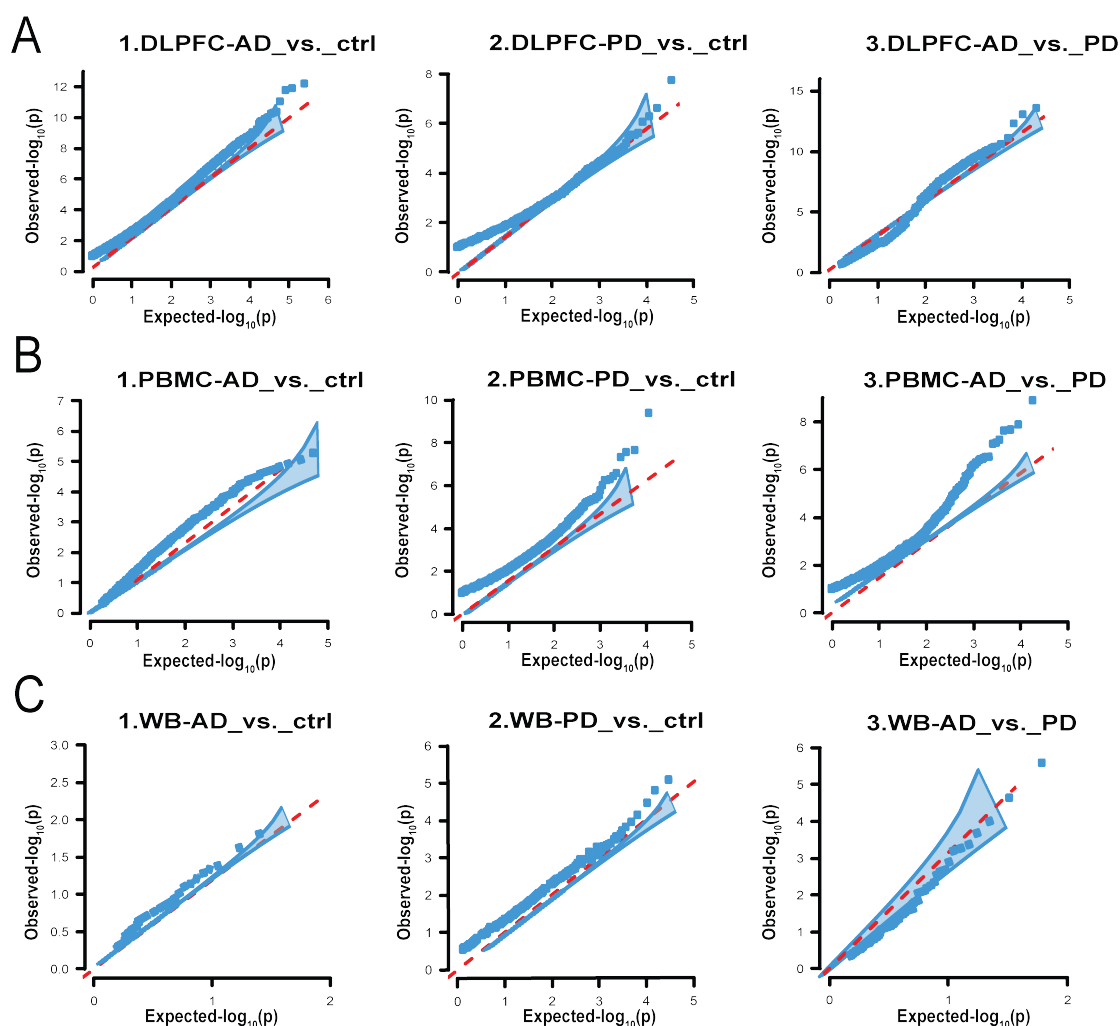


Figure 3.4 Quantile-Quantile plots (QQ-Plots) of comparing genome-wide association study (GWAS) groups.

Quantile-Quantile plots (QQ-Plots) of calculated P -values for the discovery phase of comparing genome-wide association study (GWAS) groups (1) AD-vs.-ctrl, (2) PD-vs.-ctrl, and (3) AD-vs.-PD in different tissues (A) DLPFC/BA9, (B) PBMC, and (C) WB. The observed $-\log_{10}$ of p -values ($-\log_{10}(p)$) for associated SNPs in discovery phase (Y-axis) are plotted against expected $-\log_{10}(p)$ (X-axis) under the null hypothesis (red line).

There were 252,730 SNPs associated with AD DLPFC/BA9 samples (n=133) compared to controls (n=60). PLINK clumping procedure was applied to define potential hits within the results (Table 3.1). Although all SNPs were randomly represented among all 22 chromosomes (Figure 3.5 (A1)), four SNPs exceeded genome-wide significance ($p \leq 1.0\text{E-}8$) (Figure 3.6 (A1)). The most significant signal was related to rs34956352 (chromosome (chr)4:61860228, $p \leq 6.0\text{E-}13$ and minor allele frequency (MAF)= 0.4) located within the adhesion G protein-coupled receptor L3 (*ADGRL3*). This was followed by rs1030909 (chr8:4870994, $p \leq 1.22\text{E-}12$ and MAF= 0.05) located within the CUB and sushi multiple domains 1 (*CSMD1*), rs558072 (chr10:82866406, $p \leq 1.61\text{E-}12$ and MAF= 0.2) located within the neuregulin 3 (*NRG3*), and rs7914493 (chr10:14109915, $p \leq 8.90\text{E-}12$ and MAF= 0.3) located within the FERM domain containing 4A (*FRMD4A*).

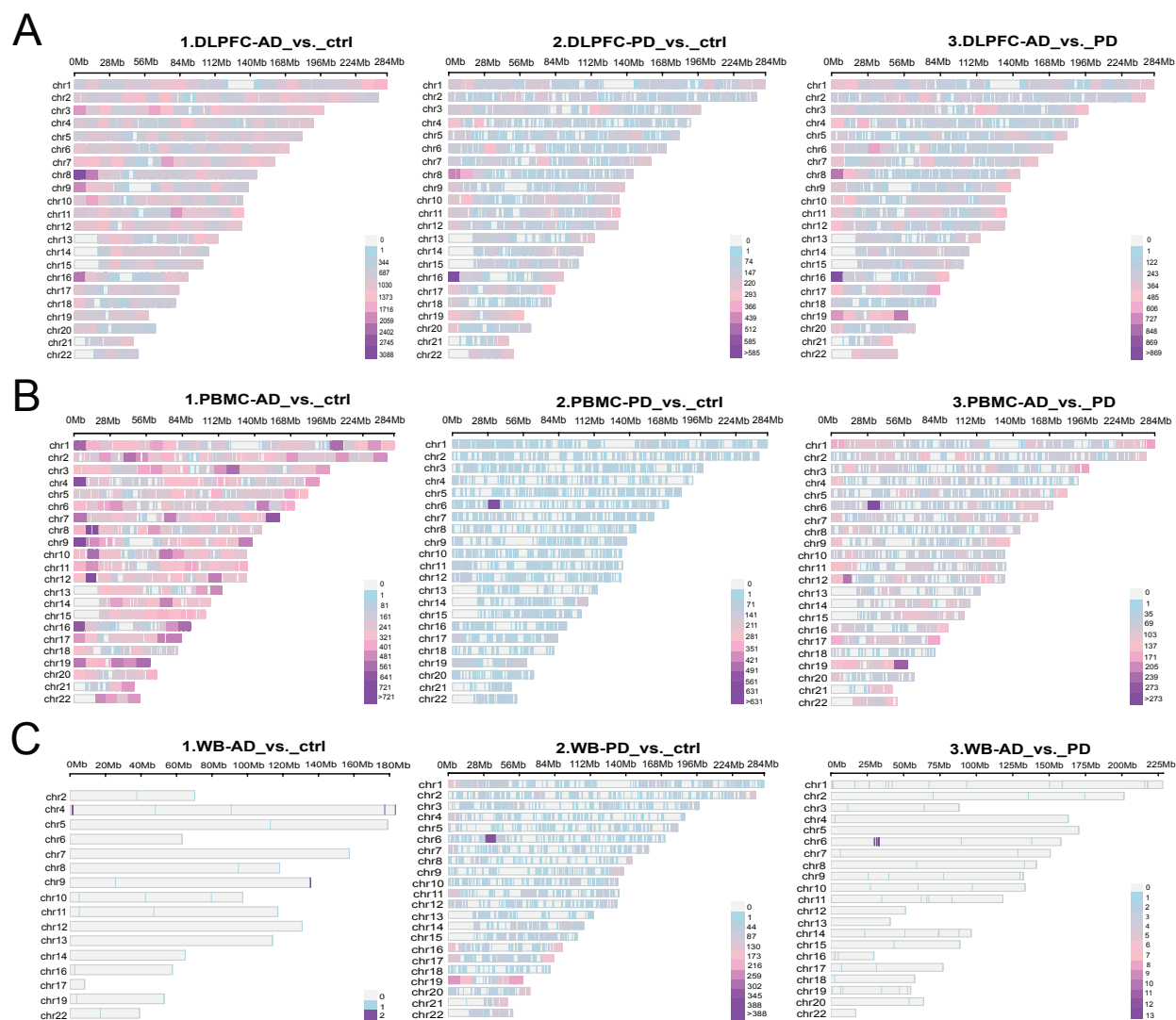


Figure 3.5 Density plots of single nucleotide polymorphism (SNP) distribution on 22 chromosomes.

SNP-density plot of single nucleotide polymorphism (SNP) distribution on 22 chromosomes of comparing GWAS groups (1) AD-vs.-ctrl, (2) PD-vs.-ctrl, and (3) AD-vs.-PD in different tissues (A) DLPFC/BA9, (B) PBMC, and (C) WB. The Y-axis shows chromosome length (Mb); the X-axis gives the chromosome number, and colors represent the number of SNP within 60kb windows.

There were 35,340 SNPs associated with PD DLPFC/BA9 samples ($n=171$) compared to controls ($n=108$). PLINK clumping procedure was applied to define potential hits within the results (Table 3.1). Among all SNPs which were randomly represented among all 22 chromosomes (Figure 3.5

(A2)), rs7329215 (chr13:67172200, $p \leq 1.0\text{E-}08$ and MAF= 0.3) located within the protocadherin 9 (*PCDH9*) was reported as significant SNP (Figure 3.6 (A2)).

There were 74,737 SNPs associated with AD DLPFC/BA9 samples (n=133) compared to PD (n=171) studies. PLINK clumping procedure was applied to define potential hits within the results (Table 3.1). All SNPs were randomly represented among all 22 chromosomes (Figure 3.5 (A3)), and three SNPs reached genome-wide significance (Figure 3.6 (A3)), including rs3021087 (chr1:630317, $p \leq 2.89\text{E-}16$ and MAF= 0.002) located near the *LOC101928626*, rs79680213 (chr11: 10508854, $p \leq 1.51\text{E-}11$ and MAF= non) located near MT-RNR2 like 8 (*MTRNR2L8*), and microRNA 4485 gene (*MIR4485*), and rs141450365 (chr6:151099698, $p \leq 7.28\text{E-}11$ and MAF= 0.0009) located near methylenetetrahydrofolate dehydrogenase (NADP⁺ dependent) 1 like gene (*MTHFD1L*).

Table 3.1 Top ten single SNP results from a genome-wide scan in terms of a smaller number of clumps of correlated SNPs per tissue-specific samples in case-control/case-case studies. Genes information related to each SNP were taken from dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

Contrast	Tissue	SNP	Chr	Position	Nearest gene	A1	A2	P-value	OR
AD-vs.-ctrl	DLPFC/ BA9	rs34956352	4	61860228	ADGRL3	G	A	6.34E-13	0.0231
		rs1030909	8	4870994	CSMD1	G	C	7.32E-13	1
		rs558072	10	82866406	NRG3	C	T	1.61E-12	0.0297
		rs7914493	10	14109915	FRMD4A	C	G	6.08E-12	0.026
		rs187731	5	72993844	FCHO2	G	T	2.95E-11	0.0151
		rs11088271	21	34088424	SLC5A3, MRPS6	G	T	3.44E-11	1
		rs201365	12	98849403	ANKS1B	A	T	3.48E-11	0.052
		rs703152	1	157008113	ARHGEF11	A	G	6.15E-11	0.0595
		rs1246194	1	111791852	KCND3	T	C	1.03E-10	1.2857
		rs7312338	12	96309295	CDK17	G	A	1.40E-10	0.0523
PD-vs.-ctrl		rs7329215	13	67172200	PCDH9	A	G	1.73E-08	0.18278
		rs1046099	14	93183156	TMEM251, MOAP1	A	G	2.33E-07	0.2084
		rs477717	1	23092768	LUZP1	G	A	6.16E-07	0.13798
		rs1394953	4	61217803	ADGRL3	C	G	2.75E-06	0.09824
		rs1469130	16	74910675	WDR59	C	T	3.41E-06	0.24309
		rs7329215	13	67172200	PCDH9	A	G	1.73E-08	0.18278
AD-vs.-PD		rs3021087	1	630317	LOC101928626	A	G	2.89E-16	1
		rs74559407	7	45252076	-	T	C	3.20E-12	1.294
		rs79680213	11	10508854	MTRNR2L8, MIR4485	A	G	1.51E-11	1.25
		rs141450365	6	151099698	MTHFD1L	C	T	7.28E-11	1.25
		rs199982713	19	12643480	-	C	T	1.19E-10	0.08332
		rs62164990	2	132254694	ANKRD30BL	T	C	5.49E-10	1
		rs10942928	5	80650052	DHFR, MTRNR2L2	G	A	7.16E-10	1.50015
		rs4087444	1	91387554	-	A	C	7.16E-10	1.50015
		rs438097	16	46806495	C16orf87	A	G	1.53E-09	1
		rs2798307	9	112581113	KIAA1958	T	G	2.56E-09	1
PD-vs.-ctrl	PBMC	rs8084	6	32443258	HLA-DRA	A	C	4.10E-10	0.04215
		rs499799	11	78079817	NDUFC2, NDUFC2-KCTD14, LOC105369401	C	G	2.28E-08	1.2222
		rs1889533	1	155978139	ARHGEF2, ARHGEF2-AS2	C	T	2.75E-08	1.25
		rs11542015	1	151993872	S100A10, LOC105371441	C	T	3.51E-07	0.04168

		rs6938946	6	35312920	DEF6	C	T	4.84E-07	23.997
		rs66966963	7	152759824	ACTR3B	C	T	5.38E-07	0.7142
		rs1143676	2	181530618	ITGA4	G	A	1.68E-06	0.22041
		rs334558	3	120094435	GSK3 β , LOC107986119	A	G	3.34E-06	5
		rs7039	3	69105192	ARL6IP5	C	G	4.12E-06	6.36268
		rs2229857	1	154601491	ADAR	T	C	4.58E-06	0.08694
AD-vs.-PD		rs202106356	4	90839056	CCSER1	G	A	1.26E-09	0.875
		rs12674366	7	133035836	CHCHD3, MIR3654	T	C	1.27E-08	34.194529
		rs1044575	2	99195607	MRPL30	G	A	2.08E-08	32.010243
		rs10445796	2	203190761	NBEAL1	A	G	5.53E-08	18
		rs2241787	19	58394796	RPS5, RNF225	T	C	7.98E-08	0.8888
		rs6715296	2	28912602	WDR43, SNORD92	C	G	3.90E-07	0.625
		rs138204380	14	53320985	LOC112268134	A	G	4.45E-07	1
		rs10634116	9	111525501	ZNF483	T	TTG	4.63E-07	1
		rs739870	20	53254310	TSHZ2	T	C	5.02E-07	1
		rs3739550	9	26892303	CAAP1	A	G	6.27E-07	26.99784
PD-vs.-ctrl	WB	rs7895609	10	113915266	NHLRC2	G	A	1.96E-06	0.4286
		rs2027585	9	93111788	CARD19	G	C	4.51E-06	0.3167
		rs3911110	2	11178351	SLC66A3	C	T	9.27E-06	0.3594
AD-vs.-PD		rs7250850	19	47082260	ZC3H4	G	C	2.41E-06	1

Abbreviations: Chr: chromosome, A1: effect allele, A2: non-effect allele, OR: odds ratio.

3.3.5 Genomic variants in peripheral blood associated with AD and PD

Risk associated SNPs were then obtained from peripheral blood tissues (PBMC and WB) to increase the possibility of early detection of AD, and PD, both by yielding a diagnostic tool and easily identifiable biomarkers.

There were 83,339 SNPs associated with AD PBMC samples (n= 22) compared to controls (n= 14). While all SNPs were randomly represented among all 22 chromosomes (Figure 3.5 (B1)), no genome-wide significant ($p \leq 5.0E-8$) signals were reported (Figure 3.6 (B1)). However, 107 variants with at least suggestively significant ($p \leq 1.0E-04$) evidence of association were observed. Additionally, PLINK clumping revealed no potential hits within the results.

There were 11,951 SNPs associated with PD PBMC samples (n= 46) compared to controls (n= 41). All SNPs were randomly represented among all 22 chromosomes (Figure 3.5 (B2)). In this study, rs8084 (chr6:32443258, $p \leq 4.0E-10$ and MAF= 0.4) located within the major histocompatibility complex, class II, DR alpha gene (*HLA-DRA*), was defined as a potential hit within the results (Table 3.1 and Figure 3.6 (B2)). This was followed by rs499799 (chr11:78079817, $p \leq 2.28E-08$ and MAF= 0.1) located near the NADH dehydrogenase [ubiquinone] 1 subunit C2 gene (*NDUFC2*),

and rs1889533 (chr1:155978139, $p \leq 2.75E-08$ and MAF= 0.2) located near the Rho/Rac guanine nucleotide exchange factor 2 gene (*ARHGEF2*).

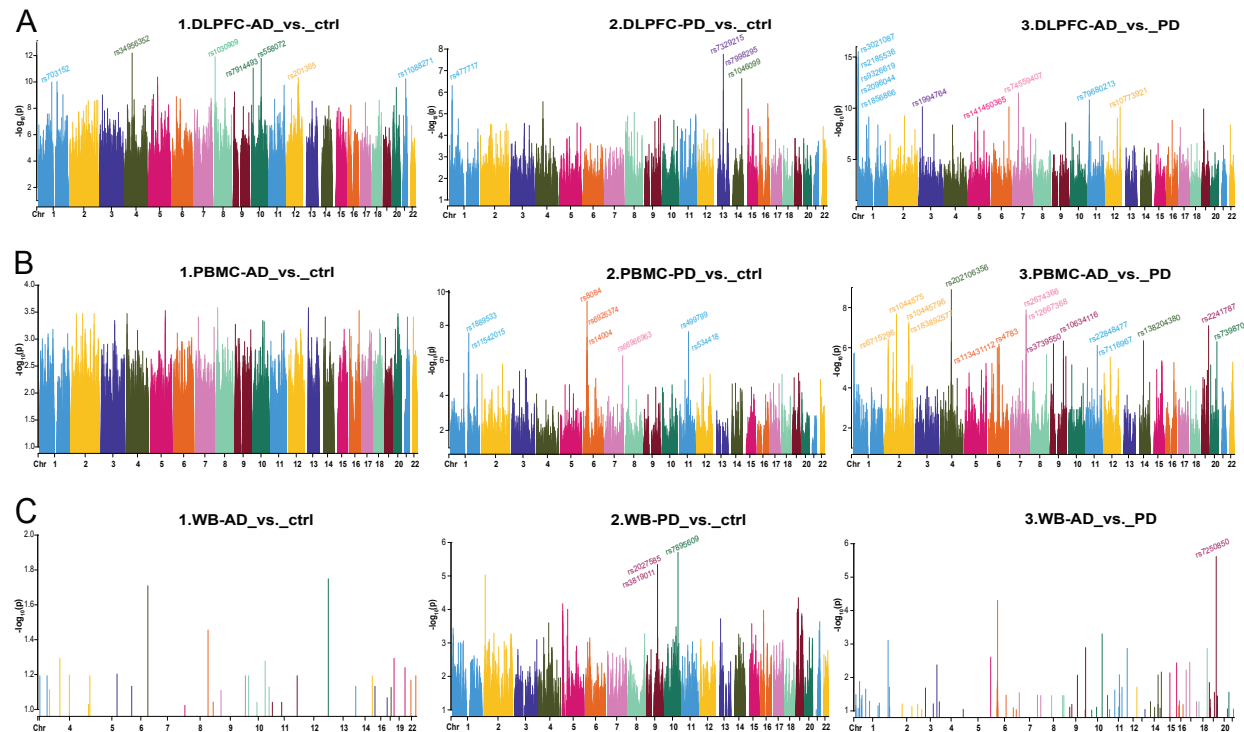


Figure 3.6 Manhattan plots summarizing the GWAS results for comparing groups.

Manhattan plots summarizing the GWAS results for comparing groups of (1) AD-vs.-ctrl, (2) PD-vs.-ctrl, and (3) AD-vs.-PD in different tissues (A) DLPFC/BA9, (B) PBMC, and (C) WB. The Y-axis is the GWAS p -value on a $-\log_{10}$ scale ($-\log_{10}(p)$); the X-axis is chromosomal position, and different colours denote significant markers on different chromosomes.

There were 18,868 SNPs associated with AD PBMC samples ($n= 22$) compared to PD ($n= 46$). All SNPs were randomly represented among all 22 chromosomes (Figure 3.5 (B3)). In this study, the lead SNP identified through PLINK clumping was rs202106356 (chr4:90839056, $p < 1.0E-9$ and MAF= 0.00004) (Table 3.1 and Figure 3.6 (B3)) located within the coiled-coil serine rich protein 1 gene (*CCSER1*). This was followed by rs12674366 (chr7:133035836, $p \leq 1.27E-08$ and MAF=1) located within the coiled-coil-helix-coiled-coil-helix domain containing 3 gene (*CHCHD3*), and

near the microRNA 3654 gene (*MIR3654*), rs1044575 (chr2:99195607, $p \leq 2.08E-08$ and MAF= 0.3) located within the mitochondrial ribosomal protein L30 gene (*MRPL30*), rs10445796 (chr2:203190761, $p \leq 5.53E-08$ and MAF= 0.4) located within the neurobeachin like 1 gene (*NBEAL1*), and rs2241787 (chr19: 58394796, $p \leq 7.98E-08$ and MAF= 0.4) located near ribosomal protein S5 gene (*RPS5*), and ring finger protein 225 gene (*RNF225*).

There were 34 SNPs reported for AD WB samples (n= 53) compared to controls (n= 32). Most of the studied SNPs were located on chromosome 4 (Figure 3.5 (C1)) and showed signals with marginal association ($p < 1.0E-02$) (Figure 3.6 (C1)). In addition, PLINK clumping revealed no independent hits within the results.

There were 10,181 SNPs associated with PD WB samples (n= 38) versus controls (n= 45). All SNPs were randomly represented among all 22 chromosomes (Figure 3.5 (C2)). In this study, three SNP variants with significant evidence of associations were observed as potential hits (Table 3.1 and Figure 3.6 (C2)). rs7895609 (chr10:113915266, $p \leq 1.97E-06$ and MAF= 0.2) located within the NHL repeat containing 2 (*NHLRC2*), rs2027585 (chr9:93111788, $p \leq 4.51E-06$ and MAF= 0.0002) located within the caspase recruitment domain family member 19 gene (*CARD19*), and rs3911110 (chr2:11178351, $p \leq 9.27E-06$ and MAF= 0.1) located within the solute carrier family 66-member 3 gene (*SLC66A3*); were among significant SNPs.

There were 94 SNPs associated with AD WB samples (n= 53) compared to PD (n= 38). Most SNPs were represented on chromosome 6 (Figure 3.5 (C3)). In this study, rs7250850 (chr19:47082260, $p \leq 2.41E-06$ and MAF= 0.3), located within the zinc finger CCCH-type containing 4 gene (*ZC3H4*) was detected as significant SNP (Table 3.1 and Figure 3.6 (C3)).

3.3.6 Functional pathways associated with AD and PD in BA9 (or DLPFC)

To explore the pathogenetic mechanism underlying AD, and PD at a systems biology level, a combination of PASCAL pipeline and WebGestaltR was used. A list of associated SNPs with the $p \leq 1.0E-01$ derived from GWAS analysis was provided as PASCAL input to calculate loci gene-based P -values. The derived gene sets were subsequently analysed for enrichment of kyoto encyclopedia of genes and genome (KEGG) pathway using WebGestaltR.

Next, shared associated variants among studied traits were analysed (Figure 3.7 A and Figure 3.7 C). For all shared associated variants, chromosomes-wise SNP density, and association between SNPs and all traits were plotted. This was followed by gene-based association analysis and identification of shared genes among studies (Figure 3.7 B and Figure 3.7 D). Additionally, gene-set enrichment analysis was performed to identify the associated biological pathways.

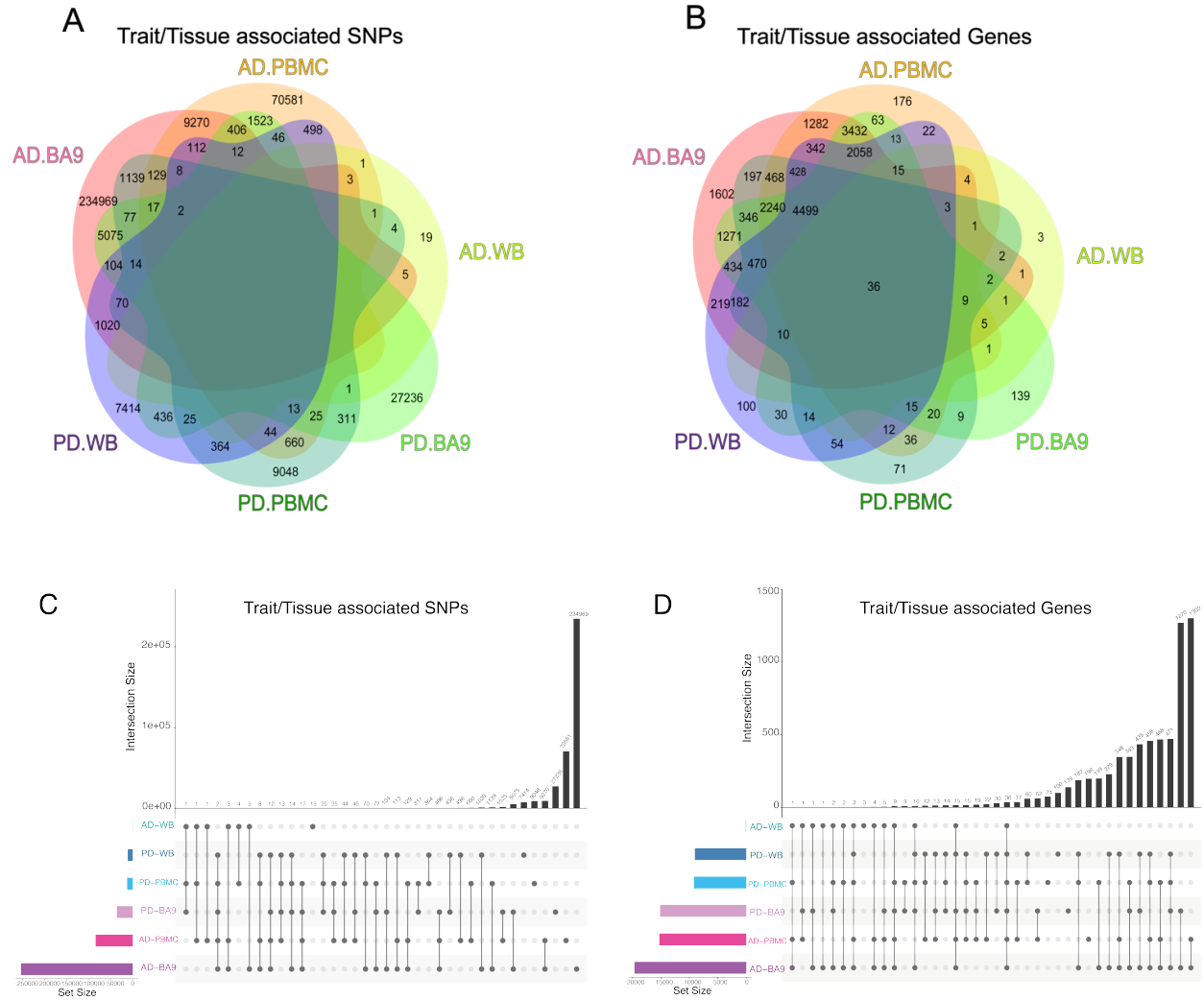


Figure 3.7 Venn diagram illustrate the overlapping single nucleotide polymorphisms (SNPs) among study groups.

(A) Venn diagram illustrate the overlapping single nucleotide polymorphisms (SNPs) between Brodmann Area 9 (BA9) of the dorsolateral prefrontal cortex (DLPFC), peripheral blood mononuclear cells (PBMC), and whole blood (WB) samples of Alzheimer's disease (AD), and Parkinson's disease (PD). In the diagram, colors represent the followings; AD.BA9 (pink), PD.BA9 (chartreuse), AD.PBMC (orange), PD.PBMC (green), AD.WB (lime), and PD.WB (purple). **(B)** Venn diagram illustrate the overlapping genes between BA9 of the DLPFC, PBMC, and WB samples of AD, and PD. In the diagram, colors represent the followings; AD.BA9 (pink), PD.BA9 (chartreuse), AD.PBMC (orange), PD.PBMC (green), AD.WB (lime), and PD.WB (purple). **(C)** Upset Plot represent the overlapping SNPs between BA9 of the DLPFC, PBMC, and WB samples of AD, and PD. In the diagram, colors represent the followings; AD.BA9 (mediumorchid), PD.BA9 (plum), AD.PBMC (maroon1), PD.PBMC (deepskyblue), AD.WB (powderblue), and PD.WB (steelblue). **(D)** Upset Plot represent the overlapping genes between

BA9 of the DLPFC, PBMC, and WB samples of AD, and PD. In the digram, colors represent the followings; AD.BA9 (mediumorchid), PD.BA9 (plum), AD.PBMC (maroon1), PD.PBMC (deepskyblue), AD.WB (powderblue), and PD.WB (steelblue).

For 19,561 associated gene in AD DLPFC/BA9 samples compared to controls, the top-ranked reported pathways were “synaptic vesicle cycle” (enrichment ratio=1.2 and $p \leq 0.000051859$) that include the genes; dynamin 1 (*DNM1*), *DNM2*, and complexin 2 (*CPLX2*) and “apelin signalling pathway” (enrichment ratio= 1.18 and $p \leq 0.0000024895$) that include the genes; myocyte-specific enhancer factor 2C (*MEF2C*), mechanistic target of rapamycin kinase (*mTOR*), mitogen-activated protein kinase 1 (*MAPK1*), and *MAPK3* (Figure 3.8 (A1), Table 3.2).

Table 3.2 *P*-values of associated genes in comparing groups per tissue.

Contrast	Tissue	Chr	Start	End	Strand	Gene	n-SNPs	P-value
AD_vs._ctrl	DLPFC/ BA9	9	130965662	131017527	+	DNM1	13	6.61E-03
		19	10823793	10944160	+	DNM2	14	3.54E-03
		5	175223609	175311023	+	CPLX2	18	5.83E-04
		5	88014057	88199922	-	MEF2C	48	5.97E-03
		22	22113946	22221970	-	MAPK1	21	4.48E-03
16		30125425	30134630	-	MAPK3	1	5.03E-02	
7		73113534	73134017	-	STX1A	1	4.56E-03	
16		31000576	31021829	-	STX1B	2	7.15E-02	
6		132778662	132834337	-	STX7	1	5.28E-02	
17		9153787	9479275	-	STX8	1	4.88E-02	
12		6571403	6579843	-	VAMP1	3	1.35E-02	
17		8062464	8066293	-	VAMP2	2	3.82E-02	
1		7831328	7841492	+	VAMP3	3	2.03E-02	
15		42787503	42825259	+	SNAP23	2	4.79E-02	
20		10199476	10288066	+	SNAP25	10	1.22E-03	
22		21213291	21245501	+	SNAP29	2	6.54E-03	
1		227916239	227968932	+	SNAP47	7	2.96E-02	
12		79257772	79845788	+	SYT1	53	1.21E-02	
1		202559724	202679551	-	SYT2	10	1.38E-02	
12		58019677	58027022	-	B4GALNT1	3	4.26E-02	
9		4490426	4587469	+	SLC1A1	5	5.50E-03	
11		35272751	35441610	-	SLC1A2	36	1.23E-03	
5		36606456	36688436	+	SLC1A3	12	6.23E-04	
2		65215578	65251000	+	SLC1A4	5	3.96E-03	
19		15060844	15121455	-	SLC1A6	1	3.50E-02	
AD_vs._PD		5	88014057	88199922	-	MEF2C	18	1.14E-02
		7	73498106	73536855	+	LIMK1	2	1.34E-02
		6	110421021	110501207	-	WASF1	4	5.70E-02
		6	151186814	151423023	+	MTHFD1L	18	1.17E-04
AD_vs._ctrl	PBMC	19	19017768	19052041	-	HOMER3	8	3.91E-02
		15	83517728	83621476	-	HOMER2	6	7.37E-03
		6	111981534	112194655	-	FYN	18	2.76E-02
		22	22113946	22221970	-	MAPK1	2	5.29E-02
		16	30125425	30134630	-	MAPK3	1	3.08E-02
		3	96533424	97467786	+	EPHA6	1	8.05E-02
		3	55499742	55521670	-	WNT5A	1	3.50E-02
		12	1683822	1756378	+	WNT5B	14	1.59E-03
PD_vs._ctrl		12	6571403	6579843	-	VAMP1	2	2.81E-02
		17	8062464	8066293	-	VAMP2	3	4.40E-02
		1	7831328	7841492	+	VAMP3	1	9.33E-02

		1	171669295	171711379	-	VAMP4	1	7.86E-02
		2	85811530	85820511	+	VAMP5	7	1.32E-02
		2	85804613	85809156	+	VAMP8	7	1.32E-02
		7	73113534	73134017	-	STX1A	2	7.08E-02
		16	31000576	31021829	-	STX1B	1	6.58E-02
		6	132778662	132834337	-	STX7	4	1.10E-02
		17	9153787	9479275	-	STX8	3	2.23E-02
		22	21213291	21245501	+	SNAP29	2	3.31E-02
		1	227916239	227968932	+	SNAP47	3	1.33E-04
		11	77779392	77791265	-	NDUFC2	3	3.65E-06
		1	155916629	155959864	-	ARHGEF2	9	3.35E-05
		1	161632904	161648444	+	FCγR2B	1	7.36E-02
AD_vs._PD		20	50003493	50179370	-	NFATC2	1	5.28E-02
		7	6414125	6443598	+	RAC1	4	3.34E-02
		11	65421066	65430443	-	RELA	2	6.65E-02
		2	99758184	99939196	+	MRPL30	29	3.40E-03
		19	58898635	58906171	+	RPS5	10	7.01E-04
AD_vs._ctrl		5	112043201	112181936	+	APC	1	6.26E-02
PD_vs._ctrl	WB	12	6571403	6579843	-	VAMP1	4	4.34E-02
		1	7831328	7841492	+	VAMP3	2	1.99E-02
		1	171669295	171711379	-	VAMP4	4	2.76E-02
		2	85811530	85820511	+	VAMP5	5	5.43E-02
		2	85804613	85809156	+	VAMP8	5	5.43E-02
		14	68117866	68141602	-	VT11B	2	2.27E-02
		1	227916239	227968932	+	SNAP47	6	1.94E-02
		7	73113534	73134017	-	STX1A	1	5.65E-02
		16	31000576	31021829	-	STX1B	2	4.51E-02
		6	132778662	132834337	-	STX7	1	6.41E-02
		17	9153787	9479275	-	STX8	1	9.28E-02
AD_vs._PD		9	102668914	102736818	+	STX17	1	7.45E-03
		6	32605182	32614839	+	HLA-DQA1	1	8.67E-02
		6	32627240	32634466	-	HLA-DQB1	1	8.67E-02
		6	32489682	32557613	-	HLA-DRB1	1	1.91E-02
		6	32485153	32557613	-	HLA-DRB5	1	1.91E-02

Abbreviations: Chr: chromosome, n-SNPs: total number of SNPs.

For 15,134 associated gene in PD DLPFC/BA9 samples versus controls, the top associated pathways that characterized were “insulin secretion” (enrichment ratio= 1.299521531 and $p \leq 1.97E-04$) that include the genes; syntaxin 1A (*STX1A*), *STX1B*, *STX7*, *STX8*, vesicle associated membrane protein 1 (*VAMP1*), *VAMP2*, *VAMP3*, synaptosome associated protein 23 (*SNAP23*), *SNAP25*, *SNAP29*, *SNAP47*, synaptotagmin 1 (*SYT1*), and *SYT2*, “metabolic pathways” (enrichment ratio= 1.1150 and $p \leq 1.3627e-9$) that include the gene; beta-1,4-N-acetyl-galactosaminyltransferase 1 (*B4GALNT1*) and “glutamatergic synapse” (enrichment ratio= 1.3 and $p \leq 9.2780e-8$) that include the genes; solute carrier family 1 member 1 (*SLC1A1*), *SLC1A2*, *SLC1A3*, *SLC1A4*, *SLC1A6*, glutamate ionotropic receptor kainate type subunit 4 (*GRIK4*), and glutamate ionotropic receptor NMDA type subunit 2D (*GRIN2D*) (Figure 3.8 (A2), Table 3.2).

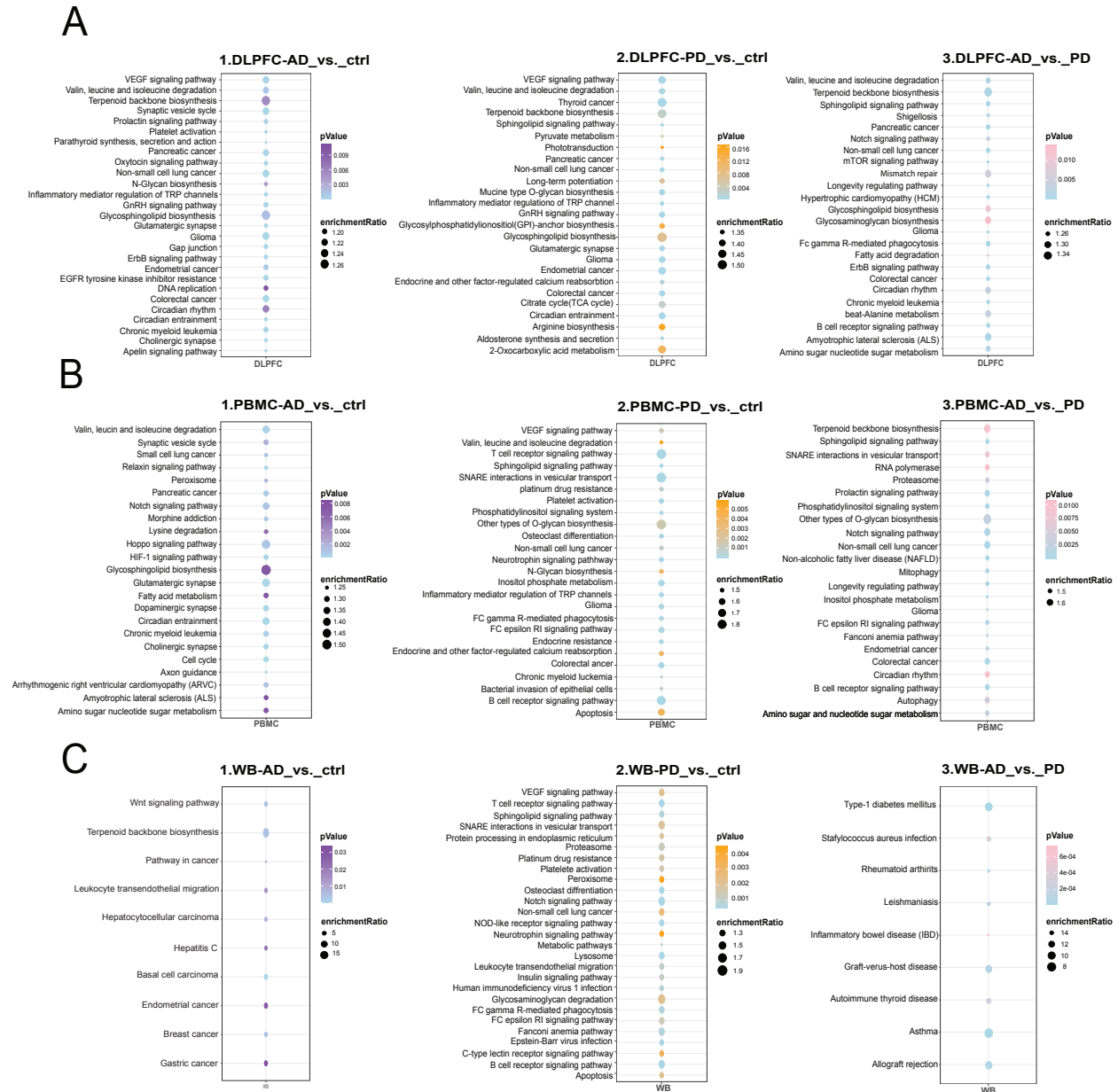


Figure 3.8 Kyoto encyclopedia of genes of genomes (KEGG) enrichment analysis for comparing groups.

Kyoto encyclopedia of genes of genomes (KEGG) enrichment analysis for comparing groups of (1) AD-vs.-ctrl, (2) PD-vs.-ctrl, and (3) AD-vs.-PD in different tissues (A) DLPFC/BA9, (B) PBMC, and (C) WB. The Y-axis represent the pathway terms with high enrichment; the X-axis represents studied tissue per phenotype. The size of the p -value (pValue) is represented by the colors of the dots. The smaller the pValue is, the closer the color is to light blue. The size of the dot indicates the enrichmentRatio of each term. Only the top 25 KEGG terms with the highest degree of enrichment were analysed.

For 17,853 associated gene in AD DLPFC/BA9 samples compared to PD, the top ranked pathways were characterized by “oxytocin signalling pathway” (enrichment ratio= 1.1779, $p \leq 0.000099069$) that include the gene; *MEF2C*, “regulation of actin cytoskeleton” (enrichment ratio= 1.1510, $p \leq 0.000096465$) that include the genes; LIM domain kinase 1 (*LIMK1*) and WAS protein family member 1 (*WASF1* or *WAVE1*) and “mTOR signalling pathway” (enrichment ratio= 1.2003, $p \leq 3.29E-06$) that include the gene; *MTHFD1L* (Figure 3.8 (A3), Table 3.2).

There was an overall of 5,707 associated SNPs simultaneously overlapping between AD and PD from DLPFC/BA9 samples ($p \leq 1.0E-01$) (Figure 3.9 A, Table 3.3). In addition, genome-wide significant loci-targeted genes showed 5,359 genes in common between AD and PD from BA9 samples (Table 3.4). The top enriched pathway was “VEGF signalling pathway” (enrichment ratio= 1.4 and $p \leq 0.0000023917$), which includes the gene; vascular endothelial growth factor A (*VEGF-A*) (Figure 3.10 A).

3.3.7 Functional pathways influencing AD and PD in peripheral blood

Pathway-based analysis was performed on peripheral blood samples (PBMC and WB) from AD, PD and NCs using SNPs with $p \leq 1.0E-01$ derived from GWAs analysis.

For 15,257 associated gene in AD PBMC compared to controls, the pathways that were specifically significant were “glutamatergic synapse” (enrichment ratio= 1.4 and $p \leq 2.2189E-9$) that include the genes; homer scaffold protein 2 (*HOMER2*), and *HOMER3* and “axon guidance” (enrichment ratio= 1.2 and $p \leq 0.000017982$) that include the genes; *FYN*, Wnt family member 5A (*WNT5A*), *WNT5B*, *MAPK1*, *MAPK3*, and EPH receptor A6 (*EPHA6*) (Figure 3.8 (B1), Table 3.2).

Overlap analysis between DLPFC/BA9 and PBMC samples in AD showed 9,953 associated variants in common (Figure 3.9 B, Table 3.3). In addition, 6,162 genes; were identified similarly in the brain and peripheral PBMC of AD (Table 3.4). The top biological pathway was “glutamatergic

synapse” (enrichment ratio= 1.3 and $p \leq 2.2629\text{E-}8$) that includes the gene; SH3 and multiple ankyrin repeat domains 3 (*SHNAK3*) and “axon guidance” (enrichment ratio= 1.2 and $p \leq 0.000062238$) that have the genes; *FYN*, glycogen synthase kinase 3 beta (*GSK3 β*), *MAPK1*, *MAPK3*, *WNT5A*, *WNT5B*, and frizzled class receptor 3 (*FZD3*) (Figure 3.10 B).

Table 3.3 *P*-values for overlapping risk loci among AD and PD in central brain and peripheral blood.

Genes information related to each SNP were taken from dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

SNP	Chr	Position	Nearest gene	C1	T1	P-value	C2	T2	P-value
rs2225201	13	67161854	PCDH9	AD	DLPFC/ BA9	9.13E-08	PD	DLPFC/ BA9	2.53E-05
rs7324857	13	67174524	PCDH9			1.63E-07			9.09E-05
rs7998295	13	67172570	PCDH9			1.84E-06			8.23E-07
rs1343	2	170082149	UBR3			1.27E-05			6.14E-05
rs2585896	13	20824722	XPO4	AD	PBMC	0.0007923	PD	PBMC	0.008509
rs2295680	14	22905226	RBM23			0.000977			0.02574
rs10002	6	151393970	-			0.001091			0.02032
rs9397400	6	151392202	ZBTB2			0.001091			0.006355
rs3216792	14	75280704	FOS			0.001109			0.0105
rs558072	10	82866406	NRG3	AD	DLPFC/ BA9	1.61E-12	AD	PBMC	0.0882
rs12513493	5	37704823	WDR70			8.71E-10			0.07364
rs9666924	11	47518455	CELF1			2.02E-09			0.01803
rs6534085	4	78938562	PAQR3			2.17E-09			0.005123
rs7207335	17	42982317	RUNDC1			3.67E-09			0.068
rs1983965	1	23093362	LUZP1	PD	DLPFC/ BA9	2.29E-06	PD	PBMC	0.002929
rs11164062	2	96185381	STARD7			3.25E-05			0.02379
rs79556420	11	10508579	MTRNR2L8, MIR4485			5.00E-05			0.004214
rs1065052	11	10534042	RNF141			0.0001151			0.0826
rs1033091	8	30065202	SARAF			0.0001203			0.002226
rs10907179	1	1319461	INTS11	PD	DLPFC/ BA9	0.0002916	PD	WB	0.005681
rs12335014	8	51820490	PCMTD1			0.0003316			0.01512
rs6682078	1	58658323	MYSM1			0.0003523			0.05281
rs10440	9	62813372	LINC01410			0.0006691			0.07697
rs187479	5	80478352	ZFYVE16			0.0007155			0.07293

Abbreviations: Chr: chromosome, C1: case number 1, C2: Case number 2, T1: tissue related to case number 1, T2: tissue related to case number 2.

For 9,175 associated gene in PD PBMC compared to controls, the most important pathways were “SNARE interactions in vesicular transport” (enrichment ratio= 1.8 and $p \leq 0.000053156$) that include the genes; *SNAP29*, *STX1A*, *STX1B*, *STX3*, *STX7*, *STX8*, *STX10*, *STX17*, vesicle transport through interaction with t-SNAREs 1B (*VTI1B*), *VAMP1*, *VAMP2*, *VAMP3*, and *VAMP8*, “metabolic pathways” (enrichment ratio= 1.085 and $p \leq 0.002969679$) that include the gene;

NDUFC2 and “Bacterial invasion of epithelial cells” (enrichment ratio= 1.575 and $p \leq 4.35E-05$) that include the gene; *ARHGEF2* (Figure 3.8 (B2), Table 3.2).

There were 485 shared SNPs between DLPFC/BA9 and peripheral PBMC in PD (Figure 3.9 C, Table 3.3). Additionally, 987 overlapping genes were detected between blood PBMC and the brain in PD (Table 3.4). The top biological pathway was “SNARE interactions in vesicular transport” (enrichment ratio= 1.7 and $p \leq 0.0017374$) that include the genes; *SNAP29*, *STX7*, *STX8*, *STX17*, *VTI1B*, *VAMP1*, *VAMP2*, and *VAMP3* (Figure 3.10 C).

For 10,941 associated gene in AD PBMC versus PD, the top associated pathways were “B cell receptor signalling pathway” (enrichment ratio= 1.4821 and $p \leq 0.00014289$) that include the genes; Fc fragment of IgG receptor IIb (*FCγR2B*), and nuclear factor of activated T cells 2 (*NFATC2*), and “sphingolipid signalling pathway” (enrichment ratio= 1.4491 and $p \leq 0.0000049764$) that includes the genes; Rho-associated coiled-coil containing protein kinase 2 (*ROCK2*), Rac family small GTPase 1 (*RAC1*), and RELA proto-oncogene, NF-κB subunit (*RELA*) and “ribosome” (enrichment ratio= 1.205 and $P \leq 0.009621496$) that include the genes; *MRPL30*, and *RPS5* (Figure 3.8 (B3), Table 3.2).

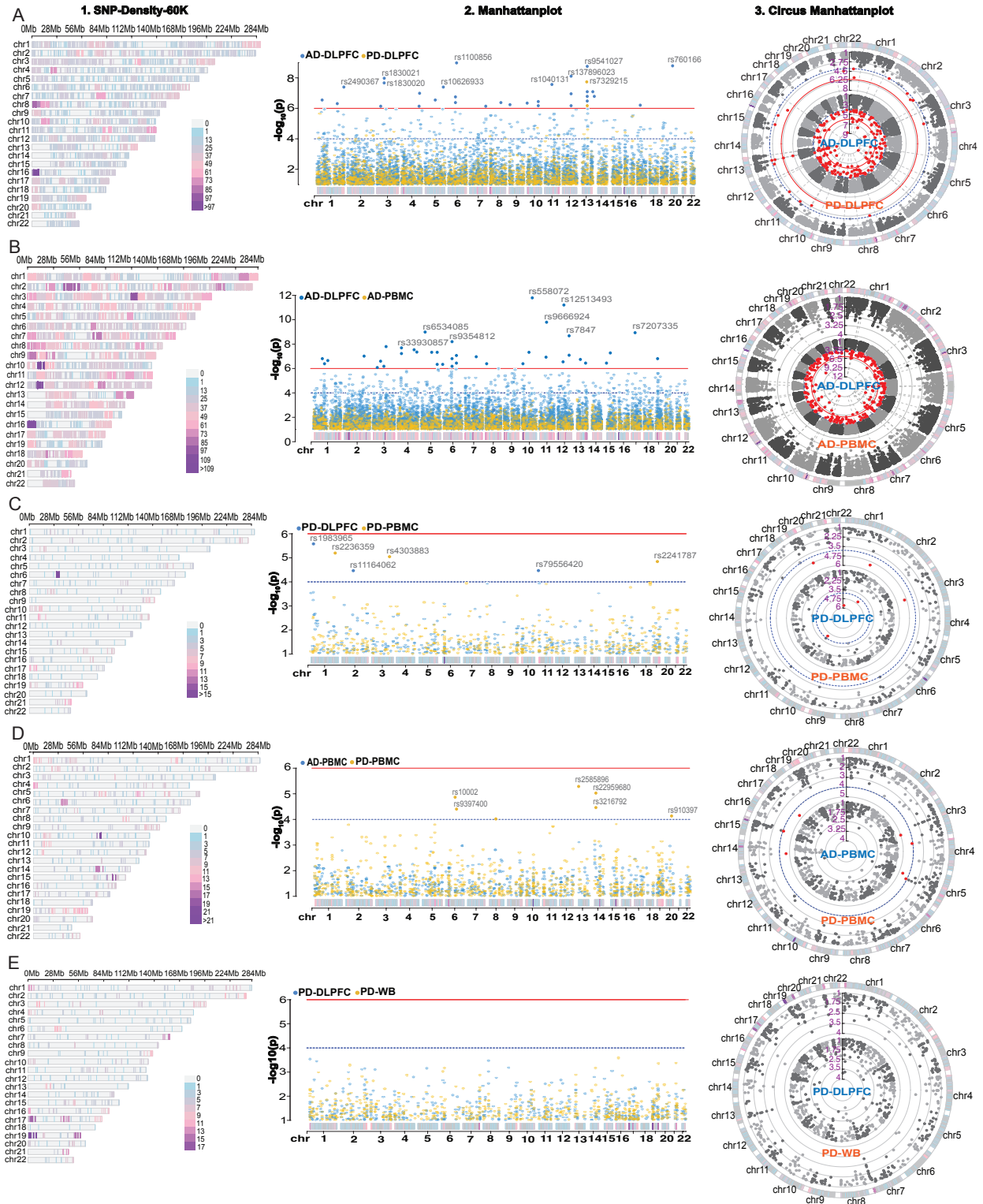


Figure 3.9 Shared genetic variants (SNPs) between comparing groups.

Investigate the shared genetic variants (SNPs) in **A**) DLPFC/BA9 samples from AD and PD, **B**) DLPFC/BA9 and PBMC samples from AD, **C**) DLPFC/BA9 and PBMC samples from PD, **D**) PBMC samples from AD and PD, and **E**) DLPFC/BA9 and WB samples from PD. From left to right, the *SNP density plot* chromosome-wise represents the number of SNPs within 60K window size. The horizontal axis shows the chromosome length (Mb); and different colour depicts SNP density. *Manhattan plot* of genome-wide association study (GWAS) results with the $-\log_{10}$ transformed p -values ($-\log_{10}(p)$) shown in chromosomal order for QC passed SNPs that were tested for a linear mixed model. The dashed lines represent the genome-wide significance threshold ($p \leq 5 \times 10^{-6}$). SNP-based GWAS *Circular Manhattan plot* displaying all associations for all outcomes plotted according to their chromosomal position on the circular line and showing the strength of the association with $-\log_{10}(p)$ on the radius. The red circular lines represent the conventional genome-wide significance threshold of $p \leq 5 \times 10^{-6}$. Phenotypes/tissues are tagged from the inner circle and outwards.

Risk-associated overlap analysis was assessed for AD and PD in PBMC samples. There were 899 associated SNPs ($p \leq 1.0 \times 10^{-1}$) in common between the two conditions (Figure 3.9 D and Table 3.3). Additionally, gene-based overlap analysis showed 1,344 shared genes between AD and PD (Table 3.4). “hippo signalling pathway” (enrichment ratio= 1.7683 and $p \leq 0.0041745$) that include the gene ajuba LIM protein (*AJUBA*), “inositol phosphate metabolism” (enrichment ratio= 1.6306 and $p \leq 0.00015828$) that include the gene; inositol polyphosphate-5-phosphatase D (*INPP5D*), “intracellular transport” (enrichment ratio= 1.2883 and $p \leq 2.2 \times 10^{-16}$) that include the genes; tubulin alpha 1a (*TUBA1A*), and *TUBA1B* and “spliceosome” (enrichment ratio= 1.3652 and $p \leq 0.0019452$) that have the genes; serine and arginine rich splicing factor 4 (*SRSF4*), *SRSF7*, *SRSF9*, splicing factor 3a subunit 2 (*SF3A2*), splicing factor 3b subunit 4 (*SF3B4*), and heterogeneous nuclear ribonucleoprotein M (*HNRNPM*); were among the corresponding top pathways (Figure 3.10 D).

For 93 associated gene in AD WB versus controls, the top associated pathway was “Wnt signalling pathway” (enrichment ratio= 6.0185 and $p \leq 0.0041195$) that include the gene; APC, WNT signalling pathway regulator (*APC*) (Figure 3.8 (C1), Table 3.2). *APC* was shown to have genetic interaction with FC γ R2A.

There were 8 SNPs in common between DLPFC/BA9 and WB tissue in AD (Table 3.3). In addition, 20 genes were identified similarly between the two tissues in AD (Table 3.4). Not shown here, the top biological pathways were “IL-17 signalling pathway” (enrichment ratio= 16.06 and $p \leq 0.060742$) and “cytokine-cytokine receptor interaction” (enrichment ratio= 5.08 and $p \leq 0.18196$) that contain the same gene; interleukin 17 receptor A (*IL-17RA*).

Table 3.4 *P*-values of overlapped genes associaed with AD and PD in central brain and peripheral blood.

Gene	Chr	Start	End	C1	T1	n-SNP	P-value	C2	T2	n-SNP	P-value
VEGF-A	chr6	43737945	43754223	AD	DLPFC/ BA9	3	3.99E-05	PD	DLPFC/ BA9	2	0.09122777
VEGF-B	chr11	64002055	64006736			6	0.00802927			1	0.07289
AQP1	chr7	30893009	30965131			2	0.01627952			2	0.0427923
AQP8	chr16	25228284	25240253			19	0.00539769			5	0.04275259
AQP10	chr1	154293591	154297801			8	0.01137732			3	0.02291243
ABCB1	chr7	87133178	87342639			8	0.00500531			1	0.03293
ABCC1	chr16	16043433	16236930			12	0.00028027			1	0.08676
GABRB1	chr4	47033294	47428447			115	7.48E-05			22	0.00206277
FAAH	chr1	46859938	46879520	AD	DLPFC/ BA9	48	0.00593514	AD	PBMIC	4	0.01655887
SHANK3	chr22	51113069	51171640			12	0.00098401			6	0.01129815
FYN	chr6	111981534	112194655			19	0.00235141			18	0.02756934
MAPK1	chr22	22113946	22221970			21	0.00434491			2	0.05292362
MAPK3	chr16	30125425	30134630			1	0.05034			1	0.03075
WNT5A	chr3	55499742	55521670			3	0.01482827			1	0.03501
WNT5B	chr12	1683822	1756378			8	6.82E-05			14	0.0015878
FZD3	chr8	28351721	28431785			12	0.02335285			9	0.01137453
GSK3 β	chr3	119540801	119813264	PD	DLPFC/ BA9	52	0.00163009	PD	PBMIC	5	0.05353029
STX1A	chr7	73113534	73134017			1	0.004561			2	0.07084222
STX7	chr6	132778662	132834337			1	0.05281			4	0.01097391
STX8	chr17	9153787	9479275			1	0.04884			3	0.0223428
STX17	chr9	102668914	102736818			2	0.04435244			3	0.04980984
VAMP1	chr12	6571403	6579843			3	0.01353486			2	0.02807905
VAMP2	chr17	8062464	8066293			2	0.03816811			3	0.0440259
VAMP3	chr1	7831328	7841492			3	0.02030515			1	0.09331
SNAP29	chr22	21213291	21245501	AD	PBMIC	2	0.00654227	PD	PBMIC	2	0.03313917
SNAP47	chr1	227916239	227968932			6	0.02736636			3	0.00013277
VTI1B	chr14	68117866	68141602			1	0.03501			1	0.09353
INPP5D	chr2	233925035	234116549			2	0.01035182			24	0.01370212
TUBA1A	chr12	49578577	49583107			3	0.05553119			1	0.07556
TUBA1B	chr12	49521566	49525304			3	0.05553119			1	0.02851
AJUBA	chr14	23440409	23451848			10	0.01630491			1	0.003659
SRSF4	chr1	29474249	29508637			3	0.03248235			9	0.00053647
SRSF7	chr2	38970740	38978636	AD	PBMIC	1	0.04042	PD	PBMIC	3	0.04075983
SRSF9	chr12	120899470	120907558			17	0.01123466			1	0.07776
SF3A2	chr19	2236815	2248678			13	0.00274745			1	0.09353
SF3B4	chr1	149895208	149900144			2	0.05866036			1	0.04042
HNRNPM	chr19	8509802	8554002			3	0.04454799			4	0.01938004
IL17RA	22	17565848	17596584	AD	DLPFC/ BA9	18	4.36E-03	AD	WB	1	6.79E-02

STX1A	chr7	73113534	73134017	PD	DLPFC/ BA9	1	0.00456	PD	WB	2	0.0708
STX7	chr6	132778662	132834337			1	0.0528			4	0.011
STX8	chr17	9153787	9479275			1	0.0488			3	0.0223
STX17	chr9	102668914	102736818			2	0.0444			3	0.0498
VAMP1	chr12	6571403	6579843			3	0.0135			2	0.0281
VAMP2	chr17	8062464	8066293			2	0.0382			3	0.044
VAMP3	chr1	7831328	7841492			3	0.0203			1	0.0933
VTIIB	chr14	68117866	68141602			1	0.035			1	0.0935

Abbreviations: Chr: chromosome, C1: case number 1, C2: case number 2, T1: tissue related to case number 1, T2: tissue related to case number 2, n-SNP: number of SNPs.

For 9,016 associated gene in PD WB compared to controls, “SNARE interactions in vesicular transport” (enrichment ratio= 1.6 and $p \leq 0.0014993$) was identified as the most prominent pathway that included the genes; *VAMP1*, *VAMP3*, *VAMP4*, *VAMP5*, *VAMP8*, *STX1A*, *STX1B*, *STX7*, *STX8*, *STX17* and *VTIIB* (Figure 3.8 (C2), Table 3.2).

There were 652 SNPs in common between DLPFC/BA9 and WB samples in PD (Figure 3.9 E, Table 3.3). In addition, 1,417 genes overlapped in both sample sets (Table 3.4), and the top biological pathways were “SNARE interactions in vesicular transport” (enrichment ratio= 1.7 and $p \leq 0.0017374$) that include the genes; *STX1A*, *STX7*, *STX8*, *STX17*, *VTIIB*, *VAMP1*, *VAMP2*, and *VAMP3* (Figure 3.10 C).

For 265 associated gene in AD WB compared to PD, “type I diabetes mellitus” (enrichment ratio= 12.4 and $p \leq 7.38E-06$) was identified as the most prominent pathway that included the genes; major histocompatibility complex, class II, DR beta 1 (*HLA-DRB1*), *HLA-DRB5*, major histocompatibility complex, class II, DQ alpha 1 (*HLA-DQAI*), and major histocompatibility complex, class II, DQ beta 1 (*HLA-DQBI*) (Figure 3.8 (C3), Table 3.2). No overlapping SNPs were found between AD and PD in WB samples.

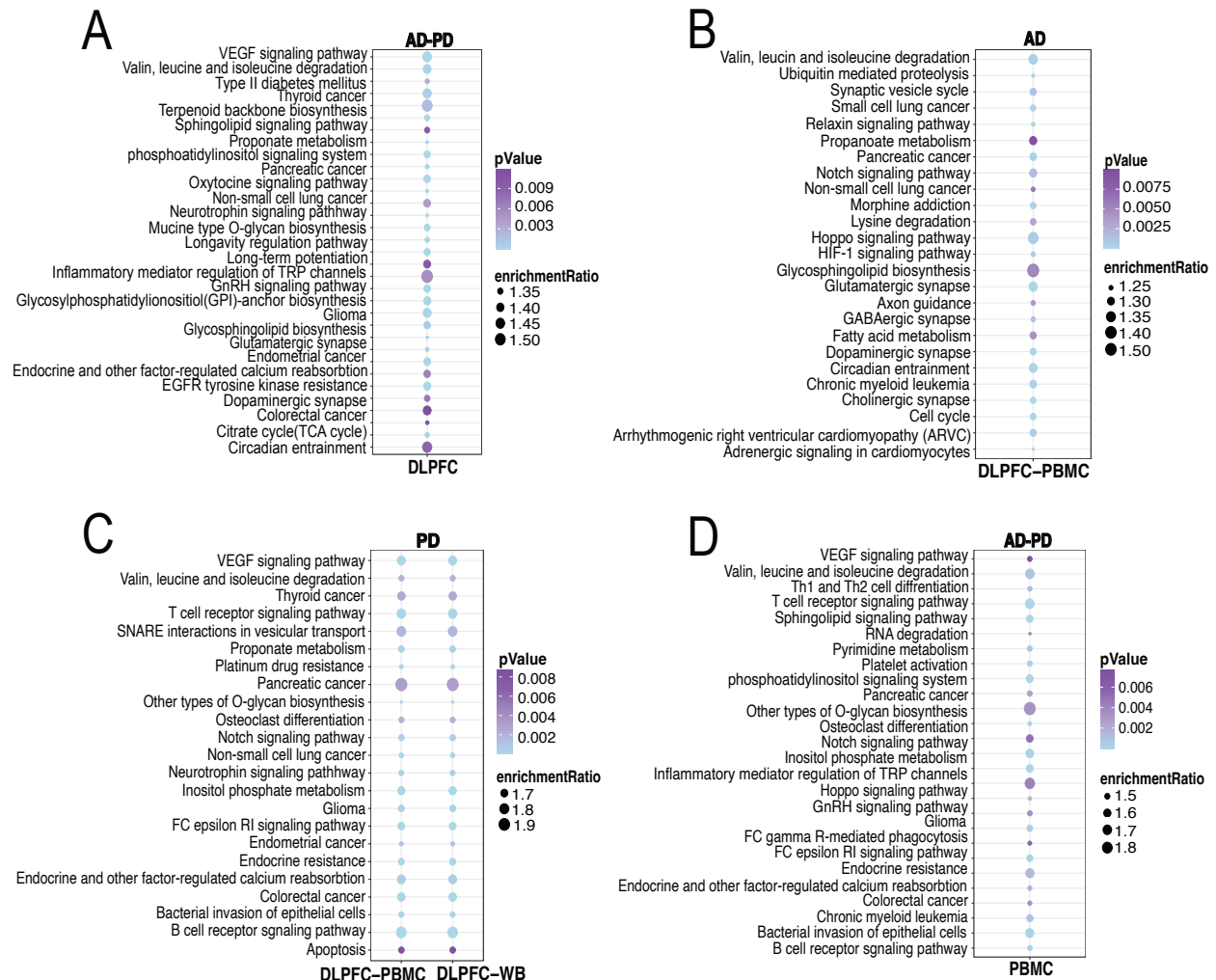


Figure 3.10 Significant overlapped kyoto encyclopedia of genes of genomes (KEGG) enriched pathways between study groups.

Significant kyoto encyclopedia of genes of genomes (KEGG) enriched pathways were plotted for overlapped SNPs-targeted genes for **(A)** DLPFC/BA9P samples between AD and PD, **(B)** For AD between DLPFC/BA9 and PBMC samples, **(C)** for PD between DLPFC/BA9, PBMC, and WB samples and **(D)** for BMC samples between AD and PD. The Y-axis represent the pathway terms with high enrichment, and the X-axis represents studied tissue per phenotype. The size of the *p*-value (*p*Value) is represented by the colors of the dots. The smaller the *p*Value is, the closer the color is to light blue. The size of the dot indicates the enrichmentRatio of each term. Only the top 25 KEGG terms with the highest degree of enrichment were analysed.

3.4 Discussions

This study explored AD and PD RNA-seq transcriptomic data obtained from BA9/or DLPFC, WB, and PBMC tissues to detect phenotype-specific common variants and their biological connection at the gene and pathway levels. The analysis at SNPs, genes and pathway levels allowed the successful identification of common related features, increasing their robustness in the association with AD and PD.

In the current study, ADGRL3 locus was identified as the major genetic risk factor for the AD-DLPFC/BA9. ADGRL3 is a well-replicated risk gene for adult attention-deficit/hyperactivity disorder (ADHD), encoding the G protein-coupled receptor latrophilin-3 (LPHN3) [215, 216]. LPHN3 has been demonstrated to play important roles in neurotransmission, glutamatergic signalling and synaptic plasticity [216]. However, the ADGRL3 locus has not been evaluated in the context of known pathways reported for AD brain here. Instead, the pathway enrichment analyses revealed *mef2c* and *DNM2* (both previously reported as LOAD-associated variants [217, 218]) as genetically associated genes with AD brain pathogenesis. The large GTPase DNM2 is a dynamin superfamily member involved in different stages of the secretory pathway. It also forms new vesicles from the Golgi network, vesicle trafficking, fusion processes and the regulation of microtubule and actin cytoskeleton dynamics [219]. Studies in the post-mortem LOAD brain showed that DNM2 down-regulation was associated with increased sequestration and scavenging of amyloid beta protein (A β) to DNM2's essential role in microtubule dynamics [217, 220]. Decreased microtubule instability might impair the capture of pre-Golgi organelles, affecting APP and BACE1 trafficking/ processing and thus impacting APP amyloidogenic cleavage [221, 222]. In the current GWAS analysis for PD-DLPFC/BA9, PCDH9 locus was reported as the significant genetic risk factor, which has been demonstrated to significantly affect signalling at neuronal synaptic junctions and long-term recognition in patients with major depressive disorder (MDD)

[223]. However, PCDH9 was not detected among significantly enriched gene sets for PD brain in the current pathway analysis. Instead, N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins (previously reported as PD-associated locus [224]), and *B4GALNT1* (previously reported as ALS-associated locus [225]) were highlighted as the most prominent genes in the pathways enriched for PD brain. *B4GALNT1*, critical for ganglioside synthesis, is responsible for the production of GM2 synthase (GM2S). This enzyme converts glucosylceramide to the simplest and most widely distributed gangliosides (GM2, GM1) that are only different due to their fatty acid residue length. Gangliosides are key components of the neuronal plasma membranes, where they participate in determining the membrane organization, neuronal function, and axon-myelin interactions and modulating protein functions [226, 227]. Pathologically, studies in the GM2-synthase knockout (KO) mice showed a reduction in gangliosides GM1 that could drive trophic signalling failure and α -syn accumulation, leading to glia/microglial activation, oxidative stress, mitochondria dysfunction and excitotoxicity, resulting in the onset of cell death, representing an accurate recapitulation of the human sporadic PD condition [228]. Additionally, accumulation of GM2 in the brain of PD mice models were also reported to potentiate classical microglial /macrophage system that led to up-regulation of pro-inflammatory cytokines (Ccl2, Ccl3, Ccl4, and Cxcl10) and down-regulation of anti-inflammatory cytokines (IL10, IL13, IL11, and IL24). This inflammatory cascade results in a loss of neurons and oligodendrocytes associated with deficiencies in memory, spatial and cognitive learning [229, 230]. Therefore, both up-regulation and down-regulation of GM2 were associated with PD pathogenesis.

Current association study derived from AD-DLPFC/BA9 compared to PD predicted noncoding RNA LOC101928626 locus to significantly differentiation AD brain from PD. However, LOC101928626 gene was not evaluated in the context of known pathways for AD, instead mef2c,

and WASF1/or WAVE1 (previously reported as MDD-associated loci [231]), a member of the WASF/WAVE family proteins highly expressed in the brain, were flagged as brain-based differentiating markers between AD and PD. WASF1 is a core component of the actin-related protein 2/3 (Arp2/3) complex known to play important roles in regulating cytoskeletal actin filaments (F-actin) dynamics required for controlling a range of cellular events such as wound healing, immune defence, embryonic development, and neuronal outgrowth [232, 233]. In AD mice models, expression of WASF1 was significantly down-regulated and observed to inhibit the trafficking of amyloid precursor protein (APP)-containing vesicles from the Golgi apparatus to the cell surface and endosomal membranes, resulting in reduced production of A β [234]. It is suggested that APP intracellular domain (AICD) binds to the WASF1 promoter, negatively regulates its transcription and downregulates WASF1 mRNA and protein expression. During disease progression, over-expression of AICD would down regulate WASF1 negative feedback and homeostatic control for A β [234, 235]. Moreover, in AD mice models, loss of WASF1 was associated with severe defects in dendritic spine morphogenesis, synapse architecture and plasticity, central nervous system myelination, and memory and cognition [235].

An important finding here is VEGF-A (previously reported as MDD-associated variant [236]), a member of the VEGF family, which was significantly enriched in the brain of both AD and PD. VEGF-A is an important signalling protein that plays a critical role in neuronal and vascular processes and is heavily involved in angiogenic regulation, neurogenesis, neuronal survival, blood-brain-barrier (BBB) permeability, and antioxidant activity [237]. Studies in PD mouse model showed that accumulation of oligomeric α -syn in the brain led to aberrant activation of astrocytes and increased release of astrocyte-released VEGF-A. Astrocyte-derived VEGF-A was reported to down-regulate tight-junction-related protein expressions such as occludin (OCLN) and claudin-5

(CLDN5), leading to an increase in BBB permeability. This was associated with induced dopaminergic (DA) neuron degeneration in the substantia nigra (SN) of PD mice [238]. Similarly, in a mouse model of AD up-regulated VEGF-A signalling contributes to OCLN down-regulation associated with BBB hyper-permeability that led to the activation of local inflammatory markers in endothelial cells, recruiting leukocytes to the site of injury, increasing the incidence of stalled capillaries, and ultimately leading to cerebral blood flow reduction [239]. This recent perspective linked BBB breakdown to cognitive impairment in AD patients [239].

Although, the current GWAS analysis for AD peripheral blood revealed no evidence of significant association, the pathway enrichment analyses pointed Homer2, 3 (previously reported as bipolar (BP)-associated variant [240]) and FC γ R2A (previously reported as PD-associated variant [241]) as the most prominent genes in AD-NC PBMC and AD-NC WB comparisons, respectively. Homer family of adaptor proteins contain a conserved N-terminal enabled/vasodilator-stimulated phosphoprotein homology 1 (*EVH1*)–like a domain that displays interactions with certain proteins implicated in calcium homeostasis, long-term potentiation, memory formation and neuronal plasticity [242]. Studies in AD mouse models showed that interactions of Homer2, 3 with the *YENPTY* motif of APP led to significant reduction of A β peptide production by inhibiting the APP trafficking and β -secretases maturation. This is known to block A β peptides generation/ gathering at the cell surface [242]. Additionally, physical interactions of Homer 2, 3 with the nuclear factor of activated T cells (NFAT) showed to negatively regulate T cell activation [243]. In the brains of AD patients, altered signalling of NFAT led to transcription of numerous molecules involved in inflammation and Ca²⁺ dysregulation, correlated with increased levels of oligomeric A β _{1–42} and the severity of cognitive impairment [244]. Studies on the postmortem AD brain tissue revealed that interaction of FC γ R2A with A β oligomers on the surface of immune response cells was

important for the decrease of A β plaque load through activation of microglia recruitment and phagocytosis [245].

In addition, FYN, a tyrosine kinase belonging to the src-family-kinase (SFK), and SHANK3, a member of the highly conserved Shank/ProSAP family (previously reported as frontotemporal lobar degeneration (FTLD)-associated variants [246]), were shared between PBMC and DLPFC/BA9 samples from AD. Directly overlapping here, SHANK3 is a large cytoskeletal post-synaptic scaffolding protein containing multiple conserved motifs, including ankyrin repeats (ANK), a proline-rich cluster (PRO) and Src homology 3 (SH3), PSD95/DlgA/Zo-1 (PDZ) and sterile α -motif (SAM), which are involved in protein-protein interactions required for synapse formation, and modulation of synaptic transmission and synaptic plasticity functions [247, 248]. Interestingly, the PRO-region encompasses the binding motifs for HOMER1, which is known to suppress SHANK3 expression/ function and was prominent in AD PBMC [249]. Studies in aged and AD models showed SHANK3 down-regulation to be accompanied by a strong reduction in the activation of intracellular signalling pathways involving Akt/S6/ERK 1/2 and eEF2 along with the development of A β and p-tau pathology leading to mild cognitive impairment and AD-type dementia [248, 250].

A novel finding here, IL-17RA, a receptor mainly expressed in somatostatin-expressing neurons, microglia, and astrocytes (previously identified as a risk factor for psoriasis and ankylosing spondylitis [251]), was the common marker between DLPFC/BA9 and WB samples from AD. Studies in rat model of sepsis indicated that IL-17RA upregulation led to downstream phosphorylation of Akt/ERK 1/2 signalling pathways, which was associated with increased inflammatory response, oxidative stress, leading to higher amyloid- β deposition and neuronal apoptosis linked with cognitive impairment [252].

HLA-DRA (in PBMC) and NHLRC2 (in WB) were shown as significant blood-based loci associated with PD pathogenesis in the current GWAS analysis. Previous studies have shown the association of increased risk for developing PD with the HLA-DRA region [253, 254]. However, it is crucial to acknowledge the high degree of polymorphism within the HLA genes, which leads to a wide range of haplotypes present in the population. Although the analysis identified an association with HLA-DRA, it is essential to consider the potential impact of different haplotypes on this association, a factor which I was unable to comprehensively address in the study. In addition, NHLRC2 is a novel pathogenic variant for FINCA disease connected with alterations in phagocytosis and multivesicular transportation that led to pathological aggregation of proteins [255, 256]. Recently, altered NHLRC2 mRNA levels have been detected in PD and AD [256]. However, these two loci were not reported as top-ranked genes in the significant pathways enriched for PD. Instead, SNARE proteins such as VAMP3, STX7, STX8, and VIT1B were obtained as genes strongly enriched in pathways that have been associated with PD. Autophagosome-localised Q_{abc} -SNAREs STX7, STX8, STX17, and VTI1B and the lysosome-localised R-SNARE VAMP3 have been shown to assemble into specific complexes to promote the fusion of the auto-phagosome with lysosomes to form the mature auto-phagolysosome (PL) [257]. It is believed that α -syn may impair auto-phagosome-to-lysosome fusion by affecting one or more of these SNARE proteins, which contributes to α -syn-associated neuronal death, possibly through modulating the composition of oligomeric α -syn because oligomeric α -syn becomes more abundant with autophagy inhibition [258]. Interestingly, signals from PD blood such as Q_{abc} -SNAREs and R-SNARE along with STX1A and VAMP2 were overlapping with those from PD brain.

CCSER1 was highlighted as differentiating locus between AD-PBMC and PD in the current GWA study. Based on a new genetic study, knockdown of CCSER1 disrupted its interaction with Nudel neurodevelopment protein 1 like 1 (NDEL1), a dynein regulator, contributing to defective dynein-dependent axon transport noted in NDs [259]. However, CCSER1 gene was not evaluated in the context of known pathways for AD. Instead, NFATC2, an inflammation-associated transcription factor, (previously reported as ALS-associated variant [260]) and FC γ 2B (previously reported as AD-associated variant [261]) were highlighted as the signature gene that differentiates AD peripheral PBMC from PD. The activation of NFATC2 in response to oligomeric A β plays a central role in the downstream activation of microglial and secretion of proinflammatory cytokines such as TNF α and IL-6 [262]. In contrast, FC γ R2B act as an immune inhibitor by reducing microglia activation and cytokine levels [263]. Up-regulation of immune inhibitory FC γ R2B in AD brains was associated with A β ₁₋₄₂-induced toxicity, including neuronal cell death, synaptic dysfunction, and behavioral defects by revealing high binding affinity to A β oligomers [264]. Additionally, HLA-DRB1, and HLA-DQB1 (previously reported as LOAD-associated variants [265, 266]) were the microglia-associated immune response variants differentiating AD from PD in peripheral WB samples. HLA-DR/HLA-DQ were shown to be associated with immunocompetence and histocompatibility, which are responsible for enhanced neuroinflammatory responses and elevated risk of AD pathogenesis in the brain [266-268]. These findings indicate the links between A β neurotoxicity and involvement of microglia-mediated peripheral immune responses in AD susceptibility and progression.

INPP5D gene (previously reported as LOAD-associated variant [269]) that encodes phosphatidylinositol-3,4,5-trisphosphate-5-phosphatase 1 (SHIP1) was reported as the shared signature between AD-PBMC and PD. INPP5D/ SHIP1 participates in the regulation of microglial

gene expression, specifically by inhibiting signal transduction initiated by activation of immune cell surface receptors, including triggering receptor expressed on myeloid cells 2 (TREM2), Fc gamma receptor (FcγR) and Dectin-1. Elevated expression levels of INPP5D/ SHIP1 in the brains of LOAD patients was associated with microglia-deficient phagocytic capacity, which leads to increased Aβ deposition and neurodegeneration [270]. Not directly here, studies in PD patients revealed an increase in phosphatidylinositol levels, which further evidence that PIP metabolism localised at the plasma membrane may be altered in this condition [271]. α-syn has been shown to preferentially localised to PI(4,5)P2 containing membranes, where it appears to inhibit phospholipase C (PLC) enzyme activity, subsequently preventing Ca²⁺ release from the endoplasmatic reticulum (ER) and abolish ERK 1/2 signalling pathway associated with both apoptotic and necrotic cell death-inducing pathways [272, 273].

3.5 Conclusions

In conclusion, based on gene association and pathway enrichment analysis, I showed that the bulk peripheral blood tissue (PBMC and WB) may not be a good source for reflecting high-impact risk-associated variants related to AD.

Though the BBB was thought to be intact during NDs, current evidence showed that disruption of BBB may be involved in both AD and PD progression, which was associated with increased angiogenesis biomarkers (e.g., VEGF-A). Another finding here, PIP lipids were discovered as shared peripheral PBMC risk biomarkers between AD and PD. However, there are differences in PIPs molecular mechanism that contribute towards the development of neurological pathologies in both NDs.

Noteworthy result of this study was the identification of Aβ-mediated microgliosis and cytokine secretion as PBMC-related marker differentiating AD from PD. Additionally, in AD-WB studies

versus PD, microglia enriched risk genes (e.g., HLA class II haplotype) were highlighted. This was followed by highlighting Homer 2,3 and FCR γ 2A as prominent features in the AD-NC PBMC and AD-NC WB comparisons, respectively. The direct interaction of these genes with amyloid peptide, were shown to be potential for modulating T cell activation, inflammatory responses, and microglial phagocytosis. As previously described in Chapter 2 these findings support the idea that anomalies in peripheral immune response may function as a mediator in AD pathogenesis.

Interestingly, instability of microtubule and actin cytoskeleton compounds (e.g., DNM2 and WASF1) associated with aggregation of A β peptides were established as pathogenic events in AD brains. Furthermore, as reported in previous Chapter, MEF2C with pivotal roles in neuronal system was also among brain-related risk factors associated with AD. Molecular activity of these three genes were shown to be regulated by cyclin-dependent kinase 5 (Cdk5) [274-276], a unique member of the cyclin-dependent kinases (Cdks); thus, making Cdk5 a highly possible candidate for development of therapeutics against AD.

In comparison, α -syn aggregation and changes in neuronal homeostasis such as altered production of SNARE proteins and gangliosides enriched in presynaptic plasma membrane (e.g., GM1, and GM2) were reported as pathologically confirmed signatures in PD brain. The important finding here, genes associated with the SNARE complex were reported as common biomarkers in the central brain and peripheral blood (PBMC and WB) of PD. This outcome is in complete consistency with the previous result reported in Chapter 2 that showed the role of SNARE complex formation defects in the pathogenesis of synucleinopathies in PD.

Little overlap here, the peripheral blood (PBMC and WB) and central brain of AD were similarly characterized with ERK 1/2 cascade known to be activate by Fyn, SHANK3, and IL-17RA. ERK 1/2 has been shown to have roles in A β and tau pathology by activating inflammatory, oxidative,

and apoptotic pathways in the CNS. This finding suggested that the ERK 1/2 pathway might be a significant tool and a valid therapeutic target in the future studies related to AD pathology.

Overall, despite the limitations of calling genomic variants from RNA-seq data, current work demonstrates the feasibility of SNP calling from RNA-seq data with moderate precision and sensitivity. The framework described in this work, comprehensively identified loci shown previously to be associated with neurodevelopmental and neurodegenerative disorders, providing support for the utility of our study design.

The selected loci and pathway enrichment analysis pointed to genetic overlaps between AD and PD centrally and peripherally. Although not significant here, results shown blood-based biomarkers reflecting pathological brain- and disease-related proteins in AD. Interesting finding here, results for PD were replicated from our previous research, indicating that PD blood might be a potential source of biomarkers related to alterations in genetic variants and gene expressions for diagnosing or tracking the initiation or progression in PD.

3.6 Potential limitations and future prospective

The current analyses have several limitations. Firstly, the sample size used for the present study is comparatively small for “GWAS standards” and was well under 50 in some instances. Accordingly, the statistical power of these analyses was low. Second, in addition to resulting in low power, small sample sizes also increase the possibility of false-positive findings, especially for infrequent variants (i.e., those with an MAF < 1-5%). Thus, independent replication ideally in larger datasets is needed to confirm the main findings of our GWAS. Third, the current GWAS have identified hundreds of susceptibility loci for complex traits (AD and PD) in the blood and central brain; however, the causal mechanisms are not yet fully understood. In the future work, the downstream effects of associated variants on molecular phenotypes (gene expression, alternative

splicing events) will be studied through the quantitative trait loci (QTL) analysis. This could serve as a powerful source for the biological interpretation of trait-associated loci, and their phenotypic consequences on the pathology of AD and PD. Fourth, a significant limitation in the analysis is the high degree of polymorphism within the HLA genes, which results in a wide range of haplotypes in the population. While I identified an association with HLA-DRA, I was unable to comprehensively analyze the impact of different haplotypes on this association. Further research is necessary to understand the role of HLA haplotype diversity in PD pathogenesis. Finally, as users of public data sets a fundamental limitation here is the lack of detailed covariate data across datasets for exposures such as smoking, pesticide exposure, alcohol and coffee intake for model adjustment and gene-environment interaction analysis. This is an important point, as the value of these datasets would increase greatly with the inclusion of well-documented environmental, disease onset, progression, and sub-phenotype data, such as diabetes status, depression, and dementia.

Chapter 4 Exploring the effect of genetic variants (SNPs) on alternative splicing and gene expression patterns in blood and brain of Alzheimer's and Parkinson's disease

4.1 Introduction

Genome-wide association studies (GWAS) have uncovered numerous robust associations between common variants and complex traits, such as Alzheimer's disease (AD), and Parkinson's disease (PD) [277-279]. The discovery of genetic variants that have a functional impact on transcriptome regulation, including alternative splicing and gene expression, is a potentially powerful means to facilitate more accurate and robust identification of associations between variants and the disease phenotype [280]. However, in the context of neurodegenerative diseases (NDs), there is a growing but still incomplete understanding of transcription variation patterns across the genome and the extent to which this variability is under genetic control [281]. Despite a substantial body of literature on quantitative trait loci (QTLs), enhancers, promoters, dynamic chromatin markers, and epigenetic regulation, further research is needed to elucidate the contribution of these factors to the transcriptional variability associated with NDs.

Alternative splicing (AS) is a crucial post-transcriptional regulatory mechanism that enables RNA-level complexity [282]. Through AS, specific exons may be included or excluded from the final mature messenger RNA (mRNA), allowing a single gene to generate multiple transcript and protein products with unique biological functions [282, 283]. Abnormalities in AS produce dysfunctional RNA isoforms that may have detrimental consequences including with age-related disorders, such as AD, and PD [21, 283]. For example, transcriptomic analysis of tau deficient cell lines revealed splicing defects in 226 genes, with intron retention representing approximately 50% of these defects and an additional 25% of these splicing defects being attributed to alternate usage

of first or last exons. Thus, aberrant alternative splicing may represent a major molecular mechanism underlying the pathophysiology of many tau-mediated neurological disorders [284].

Genetic variants, such as single nucleotide polymorphisms (SNPs), can alter exon inclusion or splice site usage, by creating or disrupting splice sites or other *cis* splicing regulatory elements within precursor mRNAs [285]. Such genetically regulated AS events were recently recognised as a primary link between genetic variation and disease [285]. Splicing quantitative trait loci (sQTL) analysis is a common method for discovering genotype-splicing associations. In an sQTL analysis, the splicing level of an alternative exon or splice site is treated as a quantitative trait and tested for association with genotypes across a population [286].

The expression QTL (eQTL) studies are like traditional genetic-association studies, but instead of associating genetic markers with discrete traits such as disease status, eQTL studies correlate genetic markers with quantitative gene expression (GE) levels [287]. Additionally, evidence from eQTL studies have shown that genomic regions where differences in the expression level of mRNA are associated with SNP genetic variation can modify disease risk [288]. In combination, the examination of direct effects of local genetic variants (*cis*-QTLs) can aid in the interpretation of GWAS loci by identifying direct links between SNPs and molecular phenotypes (i.e., AS, and GE) that characterise disease progression [289].

This work aimed to comprehensively detect postmortem brain and peripheral blood related QTLs (eQTL and sQTL) that can impact AD- and PD-risk associated GWAS variants, previously reported in **Chapter 3**. This will contribute to the understanding of the molecular mechanisms acting upon different genes in disease susceptibility loci, providing insights into genes relevant to the most common NDs and for the design of follow-up studies.

4.2 Materials and methods

4.2.1 RNA-based SNP calling and quality control (QC)

The data used in this study were obtained from the sequence read archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) database. A full description of each dataset, read mapping, calling SNPs from RNA-seq data, and the quality control (QC) procedures are available in **Chapter 3** (see Materials and methods).

4.2.2 Gene expression (GE) analysis

Raw RNA-seq reads were processed through a quality control and trimming approach using FastQC (v0.11.9, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and Trimmomatic (v0.38) tools [162] with default parameters to remove low-quality reads/bases and adaptor sequences. Trimmed reads were then aligned against GRCh38 using pseudo-aligner Kallisto (v0.48.0) [163] with default parameters to determine gene and transcript expression. This was followed by the usage of the “-filterByExpr” function from the edgeR R/Bioconductor package [165] to remove low expression genes/transcripts from the samples. Next, trimmed mean of M-values (TMM) normalization method from “-calcNormFactors” function in the edgeR package was applied to correct any compositional bias from the data. “-glmFit” function from edgeR was then applied to fit a negative binomial generalised log-linear model using raw read counts, an offset matrix, dispersion parameters, and design matrix. This was followed by the use of “-glmLRT” to conduct a likelihood ratio test for differential expression. The gene wise dispersion parameters were estimated by edgeR functions, including “-estimateGLMTrendedDisp” and “-estimateGLMTagwiseDisp”. Finally, Benjamini–Hochberg (BH) Q value < 0.05, and minimum fold change $|\log FC| \geq 1$ with the false-discovery rate (FDR) adjusted p -values ≤ 0.05 were used to nominate top differentially expressed genes.

4.2.3 Prediction of alternative splicing (AS) features

The relative quantity (percent splicing index [PSI]) of specific AS mechanisms was calculated using SUPPA2 (v2.3, https://github.com/AngryMaciek/snakemake_suppa2) with “-f ioe”, and “-e {SE, SS, MX, RI, FL}” parameters. SUPPA2 uses an GRCh38 annotation file and normalised read counts to calculate the occurrence of intron retention (IR), alternative first exons (AF), alternative last exons (AL), alternative 5' splice sites (A5), and alternative 3' splice sites (A3), mutually exclusive exons (MX), and skipped exons (SE) in a dataset. Differences of relative abundances ((Δ) PSI) were also reported by SUPPA2 with “-nan 0.1”, “--area 1000”, “--lower-bound 0.05”, and “-gc” parameters to quantify differential splicing across different samples. SUPPA2 has good modularity achieving higher accuracy compared to other tools, especially in the context of big datasets with lower sequencing depths [290].

4.2.4 Expression quantitative trait loci (eQTL) mapping

The MatrixEQTL (<https://github.com/andreyshebalin/MatrixEQTL>) R/Bioconductor package [291] with the additive linear model was used to conduct eQTL mapping analysis using SNP genotypes and GE transcript per million (TPM) values as inputs. All GE-SNP associations were included as *cis*-eQTLs if the distance between GE and SNPs was less than 100 kb. Association between GE-SNP pairs was considered *trans*-eQTLs if the distance between GE and SNP was beyond this 100 kb window (± 100 kb). A false-discovery rate (FDR) adjusted *p*-value $\leq 8.08\text{E-}05$, reported by MatrixEQTL, was used as the level of significance. To control potential confounding factors, such as batch labels, environmental exposures as well as intracellular fluctuations which often have relatively large effects on the measured GE levels [292] the following parameters were included in the analysis: covariates, gender, and age.

4.2.5 Identification of splicing quantitative trait loci (sQTL)

Correlation between SNP genotypes and PSI values of AS events was calculated for all contrasts per-tissue using the MatrixEQTL R/Bioconductor package with an additive linear model. The subsequent procedures were as per the eQTL analysis described above.

4.2.6 Gene enrichment and interaction analysis

Functional enrichment analysis was performed on the parent genes with significant eQTLs and sQTL (p -value $\leq 10E-8$) using WEB-based gene set analysis toolkit (WebGestaltR) [193], an R Bioconductor package, with functional enrichment (over-representation analysis (ORA)) of gene-ontology (biological process (BP)) and (Kyoto Encyclopedia of Genes and Genomes (KEGG)) pathways that apply a Fisher exact test to test for significance. This was followed by the usage of GeneMANIA [171] webserver to identify genetic interaction (GI) and co-expressed between genes.

4.3 Results

4.3.1 Identification of *cis*-expressed QTLs in central brain and peripheral blood from AD and PD

RNA-seq data from Brodmann Area 9 (BA9) of the dorsolateral prefrontal cortex (DLPFC)), whole blood (WB), and peripheral blood mononuclear cells (PBMC) derived from sequence read archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) database were analysed to identify the eQTL variants within ± 100 kb on either side of each gene. The eQTL analysis was performed using the normalised expression level of RNA-seq data and quality-controlled SNPs from the same samples (Table 4.1).

Table 4.1 Number of SNP genotypes, *cis*- and *trans*-eQTLs, and sQTLs for each group and tissue.

Contrast	Tissue	n SNPs	n QC-SNPs	n <i>cis</i> -eQTLs	n <i>trans</i> -eQTLs	n <i>cis</i> -sQTLs	n <i>trans</i> -sQTLs
AD_vs._ctrl	DLPFC/ BA9	17,673,896	2,001,311	199,053	5,232,237	13,567,199	294,832,256
PD_vs._ctrl		7,033,423	909,247	105,119	2,399,924	7,301,452	164,914,514

AD_vs_PD		13,698,338	1,748,877	142,803	3,504,477	19,204,902	418,157,637
AD_vs_ctrl	PBMC	11,759,877	1,825,881	2,638	66,989	20,560,596	460,562,668
PD_vs_ctrl		3,668,187	386,233	25,684	544,948	9,402,660	212,334,218
AD_vs_PD		9,102,045	1,401,604	16,974	379,494	37,570,073	815,880,701
AD_vs_ctrl	WB	25,925	2,944	24	634	7,198	156,680
PD_vs_ctrl		1,233,364	728,366	194,492	4,256,404	5,212,173	112,751,135
AD_vs_PD		451,275	144,208	756	19,013	817,354	17,860,952

Abbreviations: n SNPs: total number of filtered SNPs called from RNA-seq data through GATK pipeline, n QC-SNPs: number of SNPs after applying quality controlled (QC) steps using PLINK1.9, n *cis*-eQTL: number of predicted *cis*-eQTL, and n *trans*-eQTL: number of predicted *trans*-eQTL.

The QQ-plot related to each comparison showed that the distribution of the observed *p*-values of *cis*-eQTL, displayed earlier departure from the diagonal than that of *trans*-eQTL which indicated that *cis*-eQTL were easier to detect than *trans*-eQTL (Figure 4.1).

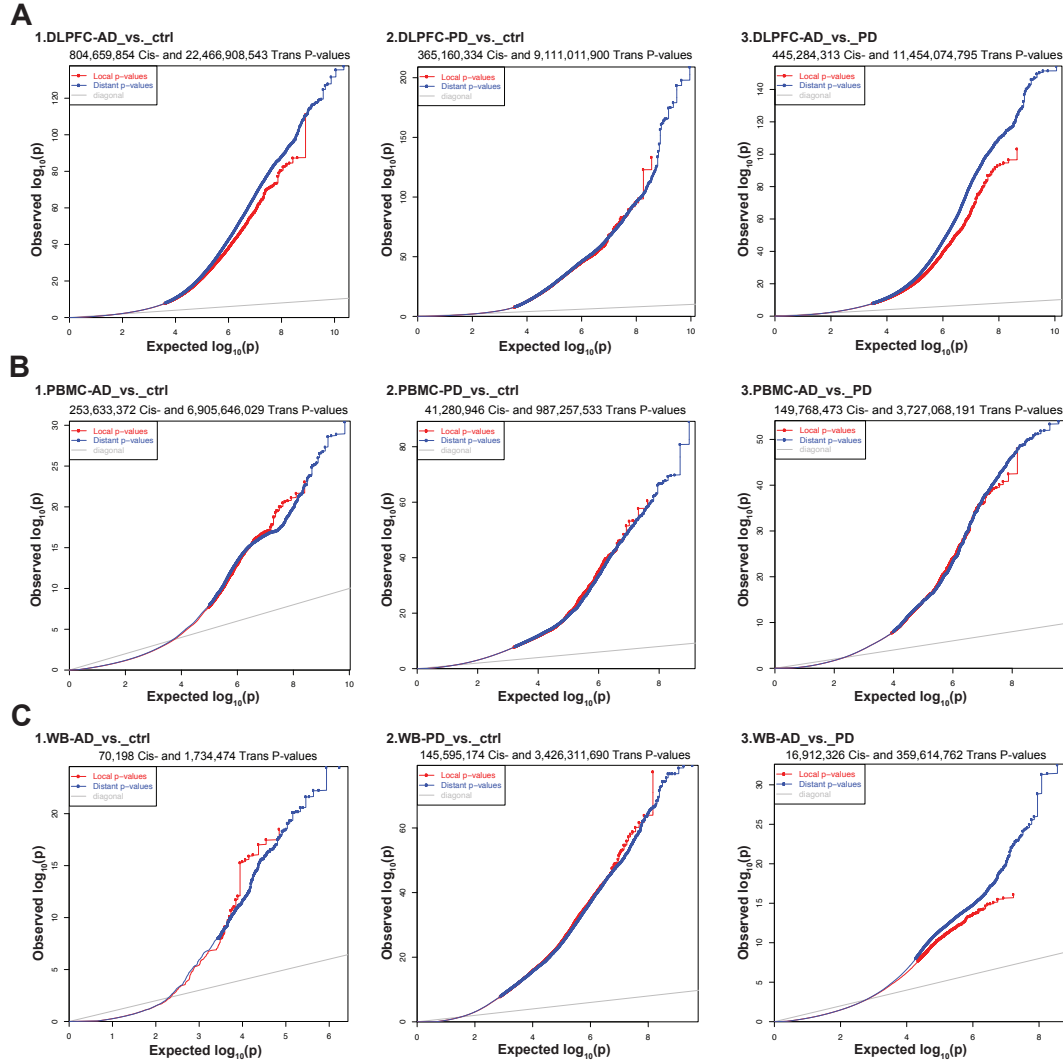


Figure 4.1 Quantile-Quantile plots (QQ-Plots) of local and distal eQTLs in AD and PD cohorts.

Quantile-Quantile plots (QQ-Plots) of calculated p -values for local and distal eQTLs of comparing groups **(1)** AD-vs.-ctrl, **(2)** PD-vs.-ctrl, and **(3)** AD- vs.-PD in different tissues **(A)** DLPFC/BA9, **(B)** PBMC and **(C)** WB. The observed $-\log_{10}$ of p -values ($-\log_{10}(p)$) for associated SNPs in discovery phase (Y-axis) are plotted against expected $-\log_{10}(p)$ (X-axis) under the diagonal hypothesis (grey line). *cis*-eQTLs, and *trans*-eQTLs are represented by red, and blue points respectively.

Next, gene enrichment analysis for *cis*-eQTL associated genes was performed to illuminate the role of *cis*-eQTL cognate genes in diverse biological processes (BPs) and KEGG pathways.

Applying an FDR adjusted p -value $\leq 8.08\text{E-}05$, 199,053 *cis*-SNP-gene pairs in the AD DLPFC/BA9 were compared to controls (Figure 4.2 A). The top enriched pathways were “synaptic vesicle cycle” (enrichment ratio= 2.4851, $p \leq 1.9470\text{e-}8$) that included: complexin 2 (*CPLX2*), and dynamin 1 (*DNMI*) and “oxytocin signalling pathway” (enrichment ratio= 1.6965, $p \leq 0.000012437$) that included: myocyte enhancer factor 2C (*MEF2C*), mitogen-activated protein kinase 1 (*MAPK1*), and *MAPK3* (Table 4.2).

There were 105,119 *cis*-SNP-gene pairs detected between SNPs and expression levels in PD-DLPFC/BA9 versus controls (Figure 4.2 B). The top enriched BPs were “synaptic vesicle cycle” (enrichment ratio= 2.6868, $p \leq 1.0502\text{e-}8$) that included: synaptosome associated protein 25 (*SNAP25*), synuclein alpha (*SNCA*), synaptotagmin 1 (*SYTI*), syntaxin 1A (*STX1A*) and “dopaminergic synapse” (enrichment ratio= 1.8574, $p \leq 0.0000073937$) that included: glutamate ionotropic receptor NMDA type subunit 2A (*GRIN2A*) and *GRIN2B*. The solute carrier family 8 member A1 (*SLC8A1*) was in common between “cardiac muscle contraction” (enrichment ratio= 2.3057, $p \leq 4.3251\text{e-}7$) and “adrenergic signalling in cardiomyocytes” (enrichment ratio= 1.5796, $p \leq 0.00090907$) pathways (Table 4.2).

There were 142,803 *cis*-SNP-gene pairs assessed associations between SNPs and gene expression levels in AD-DLPFC/BA9 samples compared to PD (Figure 4.2 C). The top enriched pathways were “oxytocin signalling pathway” (enrichment ratio= 1.8666, $p \leq 0.000044804$) that included: *MEF2C*, and myosin light chain 6 (*MYL6*), “adherens junction” (enrichment ratio= 1.8717, $p \leq 0.00419664$) that included WAS protein family member 3 (*WASF3*), and “regulation of actin cytoskeleton” (enrichment ratio= 1.4652, $p \leq 0.00501671$) that included LIM domain kinase 1 (*LIMK1*) (Table 4.2).

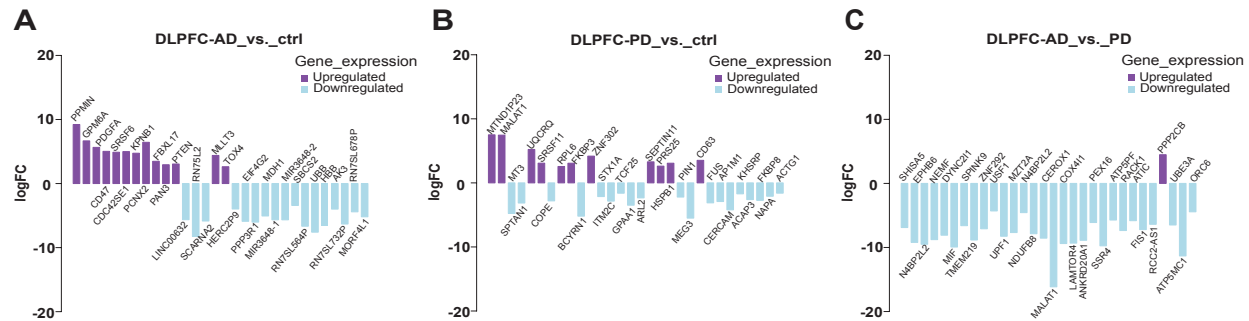


Figure 4.2 Waterfall plot of the top 30 upregulated or downregulated *cis*-eQTLs in DLPFC/BA9 samples of AD and PD cohorts.

The top 30 upregulated (purple) or downregulated (light blue) *cis*-eQTLs from DLPFC/BA9 samples in **A**) AD compared to controls, **B**) PD compared to controls, and **C**) AD compared to PD. Y-axis shows log-fold change (logFC) threshold.

The overlap analysis revealed 367 regions around *cis*-eQTLs variants in common between AD and PD in DLPFC/BA9 samples. The top enriched pathway was “retrograde endocannabinoid signalling” (enrichment ratio= 2.8913, $p \leq 0.0014474$) that included: gamma-aminobutyric acid type A receptor beta1 subunit (*GABRB1*), fatty acid amide hydrolase (*FAAH*) and *VEGF-B* (Table 4.3). Interestingly, *GABRB1* and *FAAH* genes were shown to have genetic interactions with vascular endothelial growth factor A (VEGF-A), as previously reported as common genomic biomarker between AD and PD in the brain.

There were 2,638 *cis*-SNP-gene pairs detected between SNPs and expression levels in AD-PBMC samples compared to controls (Figure 4.3 A). The top enriched pathways were “AGE-RAGE signalling pathway in diabetic complications” (enrichment ratio= 8.1562, $p \leq 0.0013714$) that included the genes; transforming growth factor beta receptor 1 (*TGFBRI*), “antigen processing and presentation” (enrichment ratio= 5.2432, $P \leq 0.055400$) that included beta-2-microglobulin (*B2M*) and “platelet activation” (enrichment ratio= 4.9235, $p \leq 0.022540$) that included Fc fragment of IgG receptor IIa (*FCγR2A*) (Table 4.2).

Table 4.2 *P*-values of *cis*-eQTL for locus identified for each contrast per tissue.

Contrast	Tissue	SNP	Nearest gene	eQTL-P-value	GWAS-P-value	GE-P-value	-logFC
AD_vs._ctrl	DLPFC/B A9	rs4073133	DNM1	3.63E-11	6.61E-03	2.68E-12	-3.813
		rs1529460	CPLX2	2.69E-11	5.83E-04	1.40E-06	-1.665
		rs30628	MEF2C	3.87E-11	5.97E-03	4.32E-05	-1.704
		rs4149444	MAPK1	4.00E-10	4.48E-03	0.013248896	-0.308
		rs73543667	MAPK3	1.99E-09	5.03E-02	0.0057106	-1.062
PD_vs._ctrl		rs226701	SNAP25	2.12E-19	1.22E-03	1.79E-10	-0.880
		rs3830694; rs71596369	SNCA	3.78E-15	2.40E-02	2.70E-09	-2.198
		rs11343990	SYT1	8.70E-15	2.89E-02	0.005397922	-0.557
		rs6968305	STX1A	4.16E-09	4.56E-03	1.19E-33	-2.138
		rs59509314; rs57519899	GRIN2A	2.35E-09	4.37E-04	-	-
AD_vs._PD		rs1053050	GRIN2B	6.47E-09	6.76E-04	-	-
		rs79050191	SLC8A1	8.80E-09	9.08E-05	0.026767807	0.664
		rs426630	MEF2C	11.79E-08	1.14E-02	6.17E-38	3.776
		rs7978034	MYL6	4.98E-21	1.39E-02	0.002371072	1.301
		rs149278597	WASF3	6.79E-11	2.85E-02	9.81E-17	1.575
AD_vs._ctrl		rs12540688	LIMK1	1.09E-10	1.34E-02	5.83E-11	1.607
		PD_vs._ctrl	rs2358996	FCγR2A	5.58E-10	3.50E-02	-
	rs882709		TGFBR1	1.42E-14	3.27E-02	-	-
	rs36100684		B2M	8.19E-09	1.73E-03	-	-
	rs4053614		STX7	5.23E-09	1.10E-02	-	-
rs1842026	STX8		5.90E-11	2.23E-02	-	-	
AD_vs._PD	rs2924135	SLC2A3	2.81E-27	6.79E-02	0.002586737	1.07	
	rs1141561	SLC44A2	1.61E-24	1.15E-03	-	-	
	rs3130617	HLA-DRB1	7.09E-18	5.78E-02	1.17E-07	2.946	
	rs3838326	HLA-DMB	7.70E-10	5.78E-02	3.49E-05	-2.042	
	rs1737087	HLA-A	4.92E-11	3.28E-02	1.01E-56	8.305	
AD_vs._ctrl	rs3093976	HLA-B	3.57E-11	8.71E-04	1.10E-14	4.823	
	PD_vs._ctrl	rs1144566	FCγR2A	2.05E-11	2.85E-02	-	-
		rs1045215	STX8	1.49E-08	9.28E-02	-	-
		rs2744430	STX11	8.68E-13	-	-	-
		rs41313373	SYTL1	1.95E-08	5.65E-02	-	-
rs2286672		VAMP1	6.33E-09	4.34E-02	-	-	
AD_vs._PD	rs2286672	VAMP2	6.33E-09	-	-	-	
	rs601071	VAPA	3.13E-14	-	-	-	
	rs2075684	HLA-DQB1	7.08E-13	8.67E-02	9.02E-09	6.998	
	rs1059517	HLA-DBP1	1.69E-08	-	6.60E-103	9.747	

Abbreviation: GE: gene expression, -logFC: log-fold change (logFC) threshold.

Overlap analysis revealed 36 regions around *cis*-eQTLs variants in common between DLPFC/BA9 and PBMC samples in AD. The top enriched pathway was “human cytomegalovirus infection” (enrichment ratio= 5.5326, $p \leq 0.015754$) that include the genes; *B2M*, and C-X-C motif chemokine ligand 12 (*CXCL12*) (Table 4.3).

In PD-PBMC samples versus controls, 25,684 *cis*-SNP-gene pairs were detected between SNPs and expression levels (Figure 4.3 B). The top enriched pathways were “immune response” (enrichment ratio= 2.3628, $p \leq 2.2\text{e-}16$) that included: *STX7* and *STX8* and “cell activation involved in immune response” (enrichment ratio= 2.8834, $p \leq 2.2\text{e-}16$) that included: *SLC2A3* and *SLC44A2* (Table 4.2).

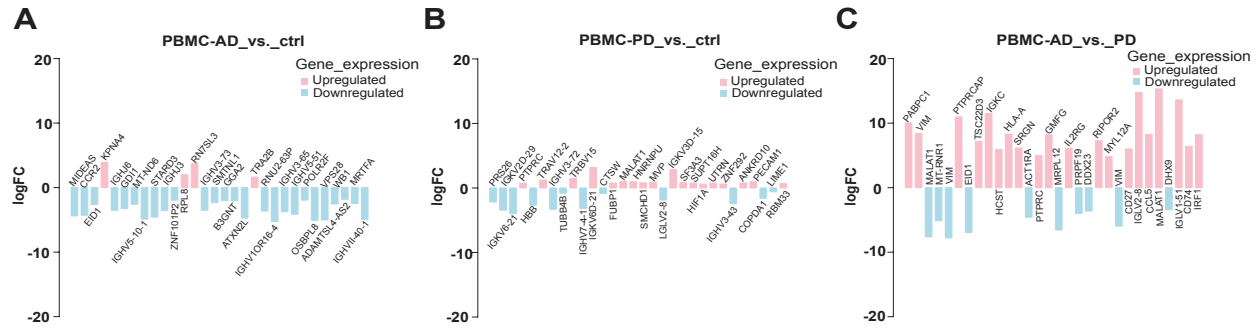


Figure 4.3 Waterfall plot of the top 30 upregulated or downregulated *cis*-eQTLs in PBMC samples of AD and PD cohorts.

The top 30 upregulated (pink) or downregulated (light blue) *cis*-eQTLs from PBMC samples in **A**) AD compared to controls, **B**) PD compared to controls, and **C**) AD compared to PD. Y-axis shows log-fold change (logFC) threshold.

In addition, overlap analysis revealed 69 regions around *cis*-eQTLs variants in common between DLPFC/BA9 and PBMC tissue from PD patients. The top enriched pathways were “protein transport” (enrichment ratio= 2.6650, $p \leq 2.2\text{e-}16$) that included *STX7* and “cytoplasmic translation” (enrichment ratio= 13.419, $p \leq 2.2\text{e-}16$) that included: eukaryotic translation initiation factor 3 subunit E (*EIF3E*), and eukaryotic translation elongation factor 2 (*EEF2*) (Table 4.3). Both genes were shown to have genetic interactions with SNARE proteins including STXBP5, STX11, and SNAP23 known to be involved in vesicle exocytosis between synapses. This pathway was shown previously to be affected in PD.

There were 16,974 *cis*-SNP-gene pairs detected between SNPs and gene expression levels in AD-PBMC samples compared to PD (Figure 4.3 C). The top enriched pathway was “antigen processing and presentation” (enrichment ratio= 4.0794, $p \leq 2.1991\text{e-}7$) that included: major histocompatibility complex (MHC) class II, DR beta 1 (*HLA-DRB1*), MHC class II, DM beta (*HLA-DMB*), MHC class I, A (*HLA-A*) and *HLA-B* (Table 4.2).

Table 4.3 *P*-values of overlapped *cis*-eQTLs between AD and PD in central brain and peripheral blood.

SNP	Nearest gene	C1	T1	eQTL-P-value	GWAS-P-value	GE-P-value	-logFC	C2	T2	eQTL-P-value	GWAS-P-value	GE-P-value	-logFC
rs1045655	GABRB1	AD	DLPFC/BA9	4.85E-10	7.87E-05	0.00593075	-0.411	PD	DLPFC/BA9	1.52E-08	2.06E-03	0.004117667	0.572
rs7518996, rs11208578	FAAH			1.24E-18	1.48E-02	-	-			8.01E-21	1.66E-02	2.08E-21	-1.464
rs1254692, rs3839946	VEGF-B			6.97E-14	8.03E-03	-	-			1.75E-13	7.29E-02	4.12E-14	-1.384
rs11294815, rs36100684	B2M	AD	DLPFC/BA9	4.02E-17	2.65E-02	4.36E-09	-4.287	AD	PBMC	8.19E-09	-	-	-
rs17293991, rs1050793	CXCL12			1.66E-29	8.41E-02	-	-			3.68E-09	-	-	-
rs143503129, rs4053614	STX7	PD	DLPFC/BA9	1.89E-13	5.28E-02	-	-	PD	PBMC	1.10E-02	1.10E-02	-	-
rs2360788, rs11776496	EIF3E			1.19E-22	2.77E-02	4.67E-12	-2.604			1.40E-22	3.72E-02	-	-
rs7247474, rs10412658	EEF2			1.02E-10	1.61E-03	6.43E-12	2.179			3.13E-13	9.35E-02	-	-
rs12817815, rs13624	TUBA1B	AD	PBMC	1.12E-08	5.55E-02	0.04180873	-1.099	PD	PBMC	1.08E-08	2.85E-02	0.049098395	0.642
rs116010141, rs2695342	ABI2	AD	DLPFC/BA9	1.59E-09	1.08E-02	2.90E-05	-1.132	AD	WB	8.02E-10	-	-	-
rs67187358, rs4652813	SNAPIN	PD	DLPFC/BA9	5.81E-18	8.98E-02	1.94E-15	-1.334	PD	WB	3.26E-12	3.31E-02	-	-
rs77151871, rs1824791	EEF2			4.97E-16	1.61E-03	6.43E-12	2.179			5.37E-20	5.61E-02	-	-
rs2360788, rs993725	EIF3E			1.19E-22	2.77E-02	0.00061902	0.963			2.27E-10	-	0.00505622	1.016

Abbreviations: C1: case number 1, C2: case number 2, T1: tissue related to case number 1, T2: tissue related to case number 2, GE: gene expression, -logFC: log-fold change (logFC) threshold.

An Overlap analysis revealed 32 regions around *cis*-eQTLs variants in common between AD and PD in PBMC samples. The top enriched pathway was “pathogenic *Escherichia coli* infection” (enrichment ratio= 0.050441, $p \leq 19.4$) that included tubulin alpha 1b (*TUBA1B*) (Table 4.3).

There were 24 *cis*-SNP-gene pairs between SNPs and gene expression levels in AD-WB samples compared to controls (Figure 4.4 A). The top enriched pathway was “staphylococcus aureus infection” (enrichment ratio= 13.338, $p \leq 0.072539$) that included Fc gamma receptor 2a (*FCγR2A*) (Table 4.2).

An overlap analysis revealed 5 regions around *cis*-eQTLs variants between DLPFC/BA9 and WB tissues from AD. The top enriched pathway was “positive regulation of Arp2/3 complex-mediated

actin nucleation” (enrichment ratio= 476.17, $p \leq 0.0020986$) that included abl interactor 2 (*ABI2*) (Table 4.3). API2 was shown to co-localised/ co-expressed with *WASF1*, an actin-nucleation promoting factor, reported previously in AD brain.

There were 19,492 *cis*-SNP-gene pairs detected between SNPs and expression levels in PD-WB samples versus controls (Figure 4.3 B). The top enriched pathway was “vesicle-mediated transport” (enrichment ratio= 2.1250, $p \leq 2.2e-16$) that included: *STX8*, *STX11*, *SYTL1*, vesicle-associated membrane protein 1 (*VAMP1*), *VAMP2*, and *VAMP* associated protein A (*VAPA*) (Table 4.2).

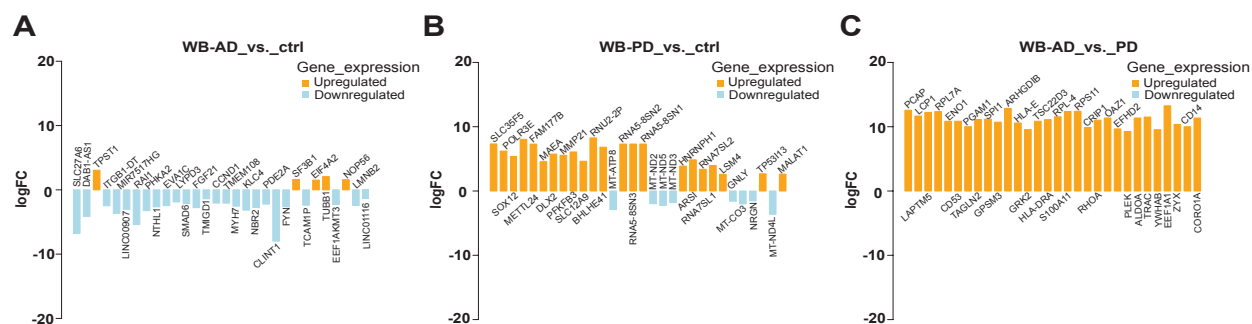


Figure 4.4 Waterfall plot of the top 30 upregulated or downregulated *cis*-eQTLs in WB samples of AD and PD cohorts.

The top 30 upregulated (orange) or downregulated (light blue) *cis*-eQTLs from WB samples in **A)** AD compared to controls, **B)** PD compared to controls, and **C)** AD compared to PD. Y-axis shows log-fold change (logFC) threshold.

An overlap analysis revealed 559 regions around *cis*-eQTLs variants between DLPFC/BA9 and WB tissues from PD. The top enriched pathways were “protein transport” (enrichment ratio= 2.4893, $p \leq 2.2e-16$) that included SNAP associated protein (*SNAPIN*) and “cytoplasmic translation” (enrichment ratio= 9.9211, $p \leq 2.2e-16$) that included *EEF2*, and *EIF3E* (Table 4.3).

Between the AD-WB and PD studies, there were 756 *cis*-SNP-gene pairs detected between SNPs and gene expression levels (Figure 4.4 C). The top enriched pathway was “Th1 and Th2 cell differentiation” (enrichment ratio= 7.6589, $p \leq 0.00045753$) that included: *HLA-DQB1*, and *HLA-*

DPBI (Table 4.2). No overlapping regions were identified around *cis*-eQTLs variants between AD and PD in WB samples.

4.3.2 Identification of *cis*-spliced QTLs in central brain and peripheral blood from AD and PD

Next, the sQTL analyses was performed (Table 4.1) to identify the number of AS events (PSI values) within the $\pm 100\text{kb}$ window for each quality-controlled SNP using linear regression. The QQ-plot related to each comparison showed that the distribution of the observed *p*-values of *cis*-sQTL, displayed earlier departure from the diagonal than that of *trans*-sQTL which indicated that *cis*-sQTL were easier to detect than *trans*-sQTL (Figure 4.5).

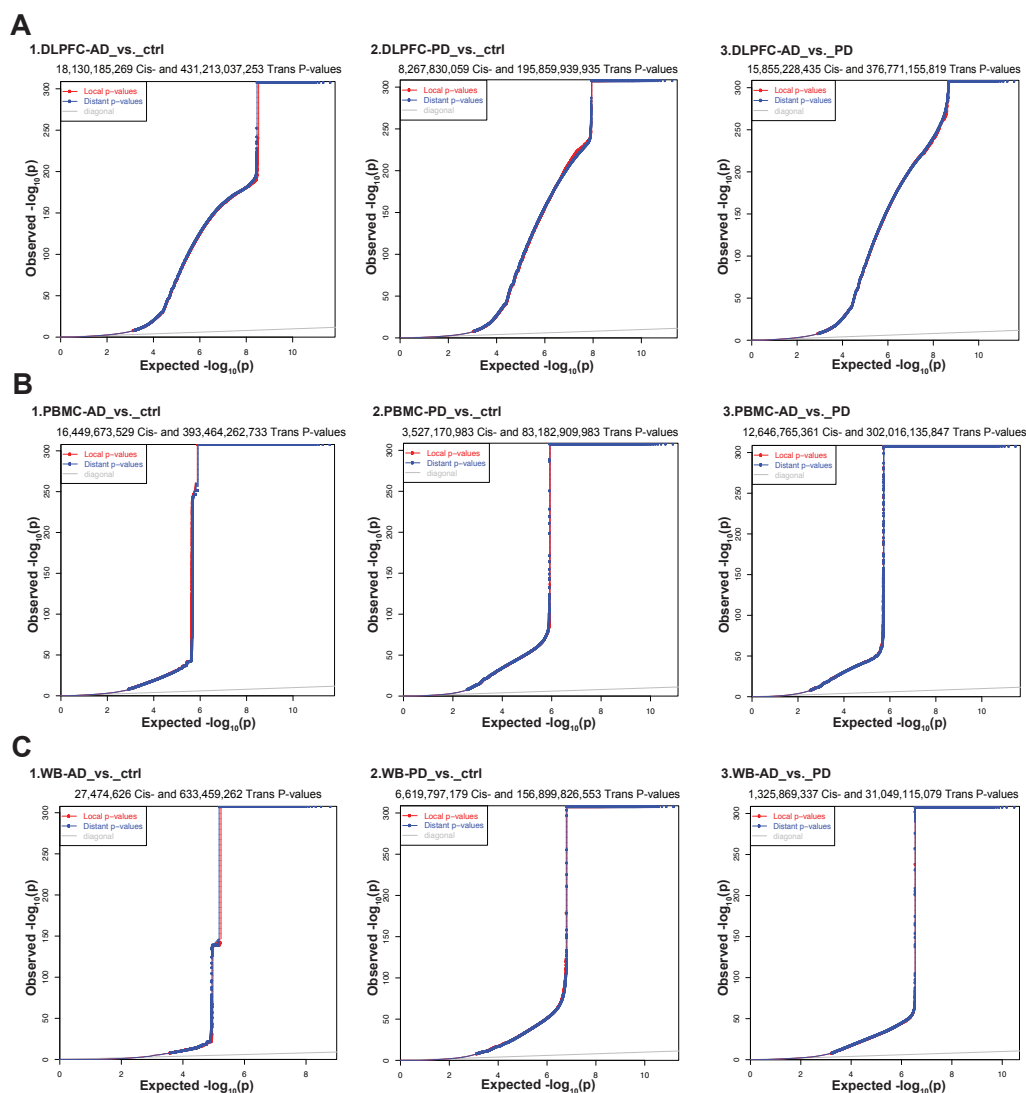


Figure 4.5 Quantile-Quantile plots (QQ-Plots) of local and distal sQTLs in AD and PD cohorts.

Quantile-Quantile plots (QQ-Plots) of calculated p -values for local and distal sQTLs of comparing groups (1) AD-vs-ctrl, (2) PD-vs-ctrl, and (3) AD-vs-PD in different tissues (A) DLPFC/BA9, (B) PBMC, and (C) WB. The observed $-\log_{10}$ of p -values ($-\log_{10}(p)$) for associated SNPs in discovery phase (Y-axis) are plotted against expected $-\log_{10}(p)$ (X-axis) under the diagonal hypothesis (grey line). *cis*-sQTLs, and *trans*-sQTLs are represented by red, and blue points respectively.

The *cis*-sQTL associated genes were then used for downstream gene enrichment analysis.

Applying the threshold of FDR recommended adjusted p -value $\leq 8.08E-05$, the sQTL analysis assessed associations between SNP and AS ratios for 13,576,242 *cis*-SNP-AS pairs in AD-DLPFC/BA9 samples compared to controls (Figure 4.6, Figure 4.7 A). The top enriched pathways were “organelle fission” (enrichment ratio= 1.4410, $p \leq 0.0000073609$) that included: dynamin 1 like (*DNM1L*), myosin XIX (*MYO19*), mitochondrial fission regulator 1 (*MTFR1*), and fission, mitochondrial 1 (*FIS1*) and “adaptive immune response” (enrichment ratio= 1.4410, $p \leq 0.0000073609$) that included: *MEF2C*, ADAM metalloproteinase domain 17 (*ADAM17*), and *SLC11A1* (Table 4.4).

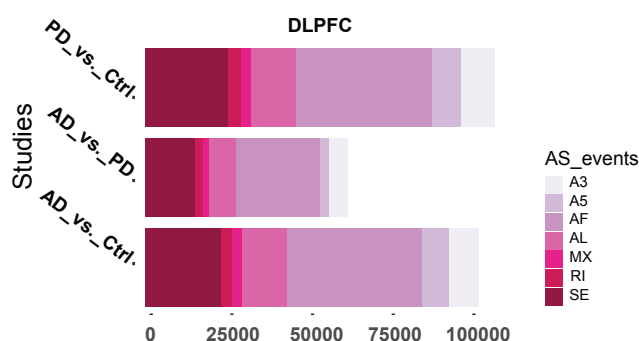


Figure 4.6 Frequency of AS events in *cis*-sQTLs identified for DLPFC/BA9 samples of AD and PD cohorts.

Number of corresponding AS events, including alternative 5'/3' splice sites (A5/A3), alternative first exon (AF), alternative last exon (AL), retained intron (RI), mutually exclusive exons (MX), and skipping exon (SE); identified in *cis*-sQTLs of comparing groups from DLPFC/BA9 samples.

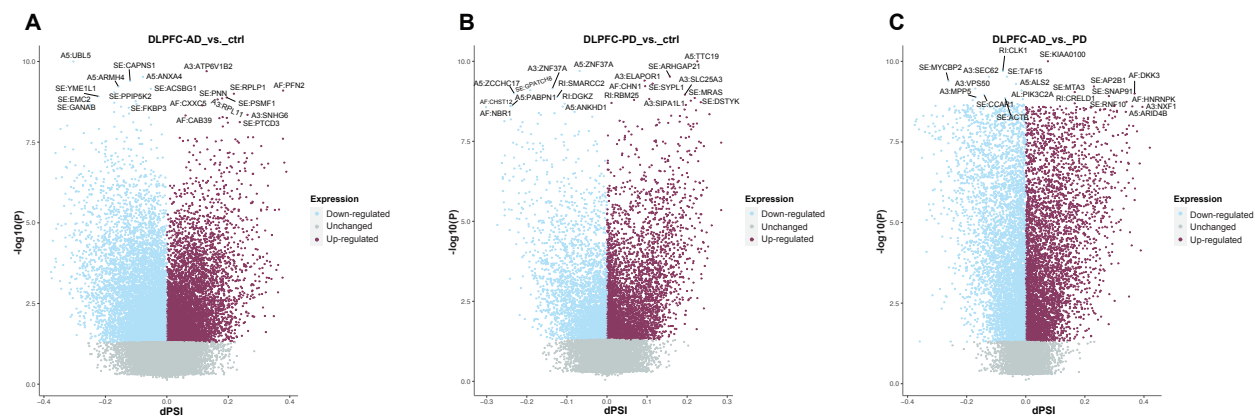


Figure 4.7 Volcano plot showing the difference of *cis*-sQTLs related AS-events in DLPFC/BA9 samples of AD and PD cohorts.

The points in the plot represent the down-regulated (light blue) and up-regulated (hot pink) sQTLs in **A**) AD compared to controls, **B**) PD compared to controls, and **C**) AD compared to PD. The observed $-\log_{10}$ of P-values ($-\log_{10}(p)$) for *cis*-sQTLs related AS-events (Y-axis) are plotted against (Δ) PSI threshold (X-axis). Genes with $-\log_{10}(p) \leq 10$ are indicated.

There were 7,301,513 *cis*-SNP-AS pairs associated with variations in splicing ratios in PD-DLPFC/BA9 versus controls (Figure 4.6, Figure 4.7 B). The top enriched pathways were “establishment of organelle localization” (enrichment ratio= 1.2627, $p \leq 0.0016727$) that included: *VAMP1*, *SYT7*, *SYT17*, *STX1A*, syntaxin binding protein 1 (*STXBPI*), *SNAP25*, and dynactin subunit 2 (*DCTN2*) and “anion transmembrane transport” (enrichment ratio= 1.4858, $p \leq 0.000054344$) that included *SLC1A2* (Table 4.4).

There were 19,204,921 *cis*-SNP-AS pairs associated with AS regulation or alteration of AS events for a specific gene in AD-DLPFC/BA9 samples compared to PD (Figure 4.6, Figure 4.7 C). The top enriched pathways were “immune response-regulating signalling pathway” (enrichment ratio=

1.4566, $p \leq 0.000027504$) that included: *MEF2C*, and *FYN* and “organelle fission” (enrichment ratio= 1.4665, $p \leq 0.000031749$) that include the gene: *WASF1* (Table 4.4).

Table 4.4 P -values of *cis*-sQTL for locus identified for each contrast per tissue.

Contrast	Tissue	SNP	Nearest gene	sQTL-P-value	GWAS-P-value	AS-event	AS-P-value
AD_vs._ctrl	DLPFC/B A9	rs2464759	DNM1L	4.01E-25	1.93E-02	AF	2.23E-05
		rs9649	MYO19	2.04E-39	7.39E-03	AL	0.000636517
		rs28659185	MTFR1	3.48E-33	2.05E-04	SE	3.75E-05
		rs201581272	FIS1	1.02E-29	2.02E-02	AF	0.001481659
		rs10939996	MEF2C	4.88E-144	5.97E-03	AF	9.00E-10
		rs9789777	ADAM17	1.42E-27	7.71E-04	SE	1.21E-05
		rs34071492	SLC11A1	7.06E-48	4.89E-03	RI	2.59E-05
PD_vs._ctrl		rs1574573	VAMP1	2.19E-16	1.35E-02	A3	0.002174425
		rs371562834, rs60924813; rs71968683	STY7	3.02E-09	3.60E-02	SE	0.000220991
		rs151226	SYT17	1.28E-22	3.51E-02	AF	1.89E-06
		rs3177697	STX1A	4.57E-10	4.56E-03	A3	1.86E-05
		rs1597813	STXBP1	2.01E-34	9.41E-03	AF	1.95E-05
		rs72732564	SNAP25	6.21E-26	1.22E-03	MX	0.014821272
		rs150249998	DCTN2	9.19E-69	2.73E-02	AL	0.002075105
		rs3847623	SLC1A2	1.53E-30	1.23E-03	AF	0.002132255
AD_vs._PD		rs3101871	MEF2C	1.51E-12	1.14E-02	AF	1.00E-10
		rs5880339	FYN	1.54E-37	4.98E-03	AF	5.00E-10
		rs4706857	WASF1	3.81E-15	5.70E-02	SE	3.27E-05
		rs10419493	HOMER3	2.10E-244	3.91E-02	AF	0.03669114
AD_vs._ctrl	PBMC	rs10419493	FYN	3.29E-17	2.76E-02	AF	0.065027258
		rs2570962	TGFBR1	2.10E-244	-	-	-
		rs55653159	B2M	6.69E-11	-	-	-
PD_vs._ctrl		rs6971864	STX1A	2.20E-13	.08E-02	SE	0.042976413
		rs2189338	STX8	2.24E-71	2.23E-02	AF	0.03232759
		rs6124358	STX16	3.82E-45	6.79E-02	AF	0.045178976
		rs79529970	STX17	6.19E-46	4.98E-02	SE	0.020528107
		rs7932	SYTL2	8.26E-37	3.50E-02	A3	0.007951215
		rs7907988	CACNB4	1.38E-74	8.65E-02	AF	0.013762202
AD_vs._PD		rs1057373	HLA-DMB	3.55E-09	5.78E-02	-	-
		rs713032	HLA-A	3.35E-39	3.28E-02	A3	7.29E-07
		rs4783	HLA-B	2.14E-09	8.71E-04	-	-
		rs1575108	HLA-G	6.84E-10	3.28E-02	-	-
AD_vs._ctrl		WB	rs41541515	FYN	3.57E-19	-	-
	rs1886682		FCER1G	1.33E-14	-	-	-
PD_vs._ctrl	rs10819437		STX17	2.21E-27	7.45E-03	AF	0.029428107
	rs2788861		STX18	2.89E-15	3.52E-02	AL	0.006492543
	rs1542610		STXBP4	4.84E-18	7.36E-02	SE	0.012401871
	rs592939		STXBP5	5.54E-18	2.46E-02	SE	0.072238822
	rs17797090		VAMP1	1.12E-15	4.34E-02	SE	0.007641222
AD_vs._PD	rs75469638		FYN	2.43E-23	-	A5	3.80E-09
	rs10946729		HLA-G	8.22E-34	-	-	-
	rs142280476		IL-37	5.45E-11	-	-	-

Abbreviation: AS-event: predicted alternative splicing events, including alternative 5'/3' splice sites (A5/A3), alternative first exon (AF), alternative last exon (AL), retained intron (RI), mutually exclusive exons (MX), and skipping exon (SE).

An overlap analysis revealed 12,066 regions around *cis*-sQTLs variants in common between AD and PD in DLPFC/BA9 samples. The top enriched pathway was “cellular response to stress” (enrichment ratio= 1.2038, $p \leq 2.2\text{e-}16$) that included: *VEGF-A*, and aquaporin 1 (*AQP1*) and “cell cycle” (enrichment ratio= 1.2246, $p \leq 2.2\text{e-}16$) that included ATP binding cassette subfamily B member 1 (*ABCB1*) (Table 4.5).

sQTL analysis assessed associations between SNP and AS ratios for 20,582,419 *cis*-SNP-AS pairs in AD-PBMC samples compared to controls (Figure 4.8, Figure 4.9 A). The top enriched pathways were “protein transport” (enrichment ratio= 1.1744, $p \leq 2.2\text{e-}16$) that included: homer scaffold protein 3 (*HOMER3*), and *FYN* and “cell cycle” (enrichment ratio= 1.1952, $p \leq 2.2\text{e-}16$) that included *TGFBR1* and “cell projection organization” (enrichment ratio= 1.1806, $p \leq 2.2\text{e-}16$) that include the gene: *B2M* (Table 4.4).

Overlap analysis revealed 12,300 regions around *cis*-sQTLs variants in common between DLPDFC/BA9 and PBMC tissues from AD. The top enriched pathway was “cellular protein localization” (enrichment ratio= 1.2113, $p \leq 2.2\text{e-}16$) that included: *HOMER3*, *FYN*, and SH3 and multiple ankyrin repeat domains 3 (*SHANK3*) (Table 4.5).

There were 9,405,357 *cis*-SNP-AS pairs associated with variations in splicing ratios in PD-PBMC samples versus controls (Figure 4.8, Figure 4.9 B). The top enriched pathways were “regulation of intracellular transport” (enrichment ratio= 1.5309, $p \leq 0.000017249$) that included: *STX1A*, *STX8*, *STX16*, *STX17*, and synaptotagmin like 2 (*SYTL2*) and “synapse organization” (enrichment ratio= 1.5848, $p \leq 0.0000074778$) that included calcium voltage-gated channel auxiliary subunit beta 4 (*CACNB4*) (Table 4.4).

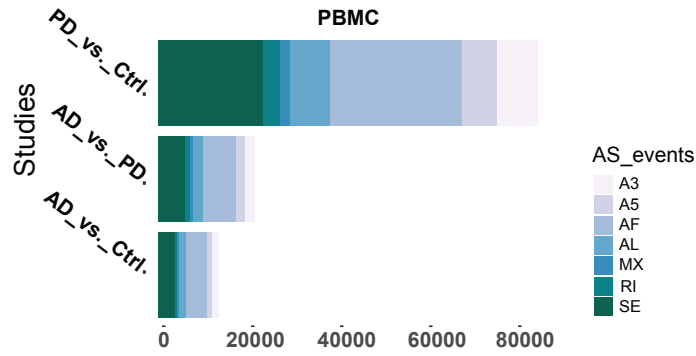


Figure 4.8 Frequency of AS events in *cis*-sQTLs identified for PBMC samples of AD and PD cohorts.

Number of corresponding AS events, including alternative 5'/3' splice sites (A5/A3), alternative first exon (AF), alternative last exon (AL), retained intron (RI), mutually exclusive exons (MX), and skipping exon (SE); identified in *cis*-sQTLs of comparing groups from PBMC samples.

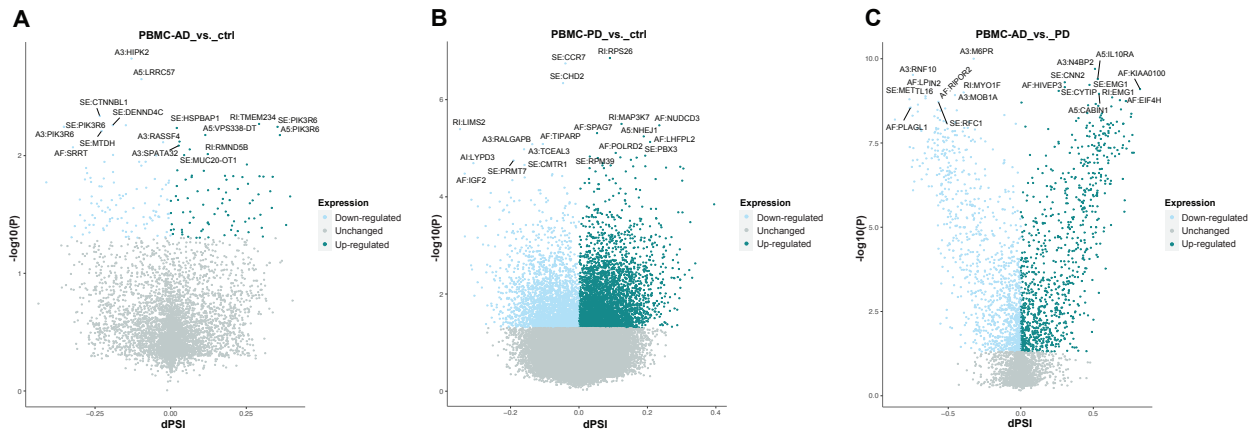


Figure 4.9 Volcano plot showing the difference of *cis*-sQTLs related AS-events in PBMC samples of AD and PD cohorts.

The points in the plot represent the down-regulated (light blue) and up-regulated (cyan) *cis*-sQTLs in **A**) AD compared to controls, **B**) PD compared to controls, and **C**) AD compared to PD. The observed $-\log_{10}$ of p -values ($-\log_{10}(p)$) for *cis*-sQTLs related AS-events (Y-axis) are plotted against (Δ) PSI threshold (X-axis). Genes with $-\log_{10}(p) \leq 10$ are indicated.

An overlap analysis revealed 10,457 regions around *cis*-sQTLs variants in common between DLPDFC/BA9 and PBMC tissues from PD. The top enriched pathways were “peptide transport”

(enrichment ratio= 1.2180, $p \leq 2.2\text{e-}16$) that included: Golgi SNAP receptor complex member 2 (*GOSR2*), *SNAP23*, *SNAP25*, *SNAP47*, *SYTL2*, *SYT9*, *STX1A*, and *STX8* and “organonitrogen compound biosynthetic process” (enrichment ratio= 1.2139, $p \leq 2.2\text{e-}16$) that included *VAPB* (Table 4.5).

Table 4.5 *P*-values of overlapped *cis*-sQTLs between AD and PD in central brain and peripheral blood.

SNP	Nearest gene	C1	T1	sQTL-P-value	GWAS-P-value	AS-event	AS-P-value	C2	T2	sQTL-P-value	GWAS-P-value	AS-event	AS-P-value
rs2840791, rs779472	VEGF-A	AD	DLPFC/BA9	1.04E-156	3.99E-05	AF	0.000260412	PD	DLPFC/BA9	8.15E-185	9.12E-02	-	-
rs3840611, rs60272141	AQP1			9.12E-14	1.63E-02	AF	0.004106927			1.54E-19	4.28E-02	AF	0.00902305
rs12668675, rs17170619	ABCB1			3.88E-11	5.01E-03	-	-			1.10E-09	3.29E-02	SE	0.076342776
rs10421518, rs1029837	HOMER3	AD	DLPFC/BA9	7.64E-11	3.48E-03	AF	0.017021069	AD	PBMC	1.72E-13	3.91E-02	AF	0.036691137
rs1766456, rs936970	FYN			1.16E-19	2.35E-03	SE	8.33E-06			43.29E-17	2.76E-02	AF	0.065027258
rs8113568, rs5765916	SHANK3			1.76E-08	9.84E-04	SE	0.000539034			1.56E-09	1.13E-02	-	-
rs72732564, rs5814358	SNAP23	PD	DLPFC/BA9	6.21E-26	4.79E-02	AF	1.01E-05	PD	PBMC	2.65E-65	-	-	-
rs6123070, rs2423368	SNAP25			2.51E-09	1.22E-03	MX	0.014821272			6.49E-29	-	A5	0.003959505
rs1042295, rs1139620	SNAP47			2.90E-15	2.96E-02	AL	0.094957472			1.54E-24	1.33E-04	AF	0.052358713
rs10565723, rs1049666	SYTIL			5.83E-10	2.87E-02	A3	0.077961222			6.02E-11	3.50E-02	SE	0.084340043
rs10565723, rs145861733	SYT9			6.38E-10	1.93E-02	AF	0.027582991			3.17E-10	-	AF	0.009107253
rs3177697, rs6971864	STX1A			4.57E-10	4.56E-03	A3	1.86E-05			2.20E-13	7.08E-02	SE	0.042976413
rs9906952, rs2189338	STX8			3.39E-11	4.88E-02	AF	0.002118381			2.24E-71	2.23E-02	AF	0.03232759
rs7474, rs242538	VAPA			3.28E-13	1.34E-03	AF	0.00083159			8.61E-14	3.23E-02	SE	0.023782921
rs1076185, rs1010228	GOSR2			1.15E-08	4.71E-02	A3	0.00083159			1.28E-10	3.05E-02	-	-
rs20033268, rs4376957	TUBA1B	AD	PBMC	1.73E-22	5.55E-02	SE	0.04286744	PD	PBMC	2.30E-58	2.85E-02	RI	0.006334101
rs10163477, rs1069908	INPP5B			5.98E-14	1.94E-02	AL	0.066841543			5.20E-10	7.06E-02	A5	0.007854989
rs55845795, rs849749	ABL2	AD	DLPFC/BA9	2.06E-72	2.59E-03	AF	1.15E-05	AD	WB	2.09E-12	-	AF	0.054113204
rs61806094, rs849749	FCRL3			5.10E-178	-	-	-			1.71E-12	-	-	-
rs1766456, rs41541515	FYN			1.16E-19	2.35E-03	SE	8.33E-06			3.57E-19	-	-	-
rs2341983, rs9879812	IL-17RD			7.93E-16	1.32E-03	-	-			1.97E-10	-	-	-
rs6485461, rs4244821	IL-18			6.94E-10	3.92E-03	-	-			1.33E-14	-	-	-
rs1184662, rs3779234	IL-6			1.10E-13	-	A3	0.011683124			6.07E-11	-	-	-
rs10565723, rs10437597	SYTL1	PD	DLPFC/BA9	2.28E-10	3.51E-02	AF	6.17E-06	PD	WB	1.05E-10	5.65E-02	A3	0.018872261
rs7974618, rs10843222	VAMP1			4.27E-10	1.35E-02	A3	0.002174425			3.82E-13	4.34E-02	SE	0.007641222
rs7974618, rs108432	VAMP3			4.27E-10	1.35E-02	A3	0.002174425			3.82E-13	4.34E-02	SE	0.007641222
rs1886682, rs3737578	FCER1G	AD	WB	1.33E-14	-	-	-	PD	WB	3.31E-11	-	-	-
rs10798666, rs608881	FCγR1B			2.09E-12	-	-	-			1.09E-23	-	-	-

rs28399248, rs1772441	FCγR3B			2.52E-13	-	-	-			2.34E-59	4.50E-02	A3	0.02611329
--------------------------	--------	--	--	----------	---	---	---	--	--	----------	----------	----	------------

Abbreviations: C1: case number 1, C2: case number 2, T1: tissue related to case number 1, T2: tissue related to case number 2, AS-event: predicted alternative splicing events, including AS-event: alternative 5'/3' splice sites (A5/A3), alternative first exon (AF), alternative last exon (AL), retained intron (RI), mutually exclusive exons (MX), and skipping exon (SE).

There were 37,582,420 *cis*-SNP-AS pairs associated with AS regulation or alteration of AS events for a specific gene in AD-PBMC samples compared to PD (Figure 4.8, Figure 4.9 C). The top enriched pathways were “protein-containing complex assembly” (enrichment ratio= 1.1734, $p \leq 2.2\text{e-}16$) that included *HLA-DMB* and “endocytosis” (enrichment ratio= 1.2202, $P \leq 0.000015827$) that included: *HLA-A*, and *HLA-B* (Table 4.4).

An overlap analysis revealed 10,844 regions around *cis*-sQTLs variants in common between AD and PD in PBMC samples. The top enriched pathways were “intracellular transport” (enrichment ratio= 1.2792, $p \leq 2.2\text{e-}16$) that included *TUBA1B* and “small molecule metabolic process” (enrichment ratio= 1.1970, $p \leq 2.2\text{e-}16$) that included inositol polyphosphate-1-phosphatase (*INPP5B*) (Table 4.5). INPP5B was shown to co-expressed with INPP5D, a previously associated variant.

sQTL analysis assessed associations between SNP and AS ratios for 7,356 *cis*-SNP-AS pairs in AD-WB samples compared to controls (Figure 4.10, Figure 4.11 A). The top enriched pathway was “protein localization to membrane” (enrichment ratio= 1.5943, $p \leq 8.3508\text{e-}9$) that included: *FYN*, and Fc fragment of IgE receptor Ig (*FCER1G*) (Table 4.4).

An overlap analysis revealed 2,821 regions around *cis*-sQTLs variants in common between DLPFC/BA9 and WB samples from AD. The top enriched pathways were “positive regulation of catalytic activity” (enrichment ratio= 1.6443, $p \leq 3.4405\text{e-}9$) that included: ABL proto-oncogene 2, non-receptor tyrosine kinase (*ABL2*), Fc receptor like 3 (*FCRL3*), and *FYN* and “protein

phosphorylation” (enrichment ratio= 1.2904, $p \leq 2.5966e-7$) that included: interleukin 17 receptor D (*IL-17RD*), *IL-18*, and *IL-6* (Table 4.5).

There were 5,212,719 *cis*-SNP-AS pairs associated with variations in splicing ratios in PD-WB samples versus controls (Figure 4.10, Figure 4.11 B). The top enriched pathways were “endosomal transport” (enrichment ratio= 1.4703, $p \leq 0.0036229$) that included: *STX17*, *STX18*, *STXBP4*, and *STXBP5* and “vesicle localization” (enrichment ratio= 1.6090, $p \leq 0.000014555$) that included *VAMP1* (Table 4.4).

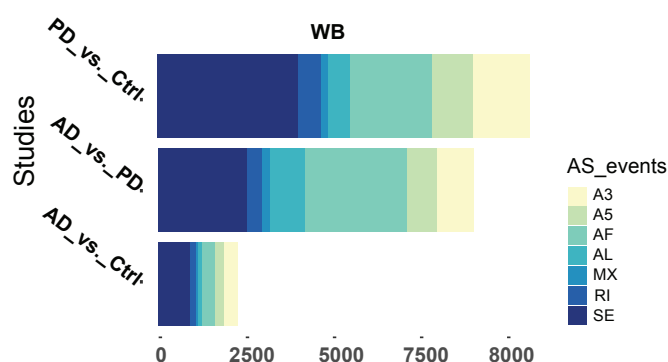


Figure 4.10 Frequency of AS events in *cis*-sQTLs identified for WB samples of AD and PD cohorts.

Number of corresponding AS events, including alternative 5'/3' splice sites (A5/A3), alternative first exon (AF), alternative last exon (AL), retained intron (RI), mutually exclusive exons (MX), and skipping exon (SE); identified in *cis*-sQTLs of comparing groups from WB samples.

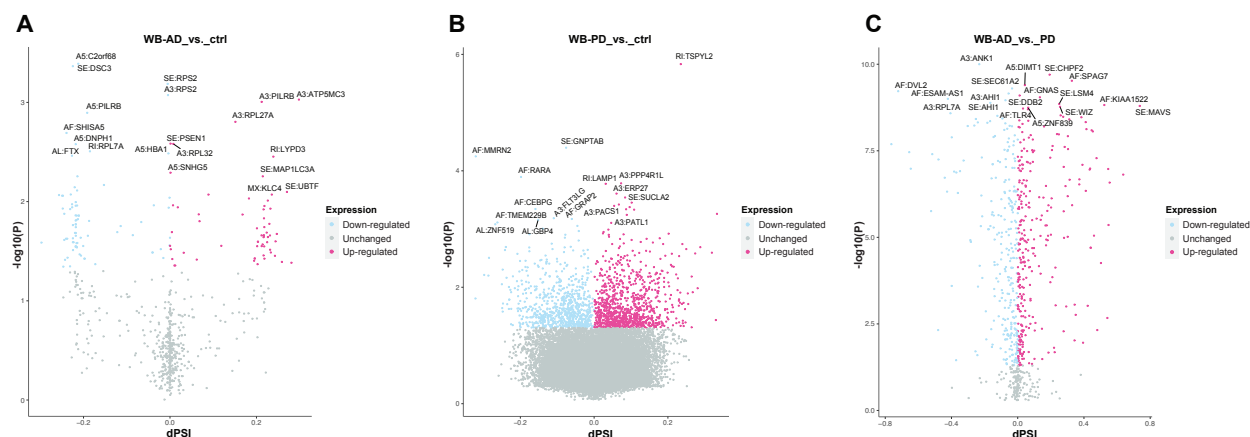


Figure 4.11 Volcano plot showing the difference of *cis*-sQTLs related AS-events in WB samples of AD and PD cohorts.

(Δ) PSI is plotted against the $-\log_{10}(p)$ value. The points in the plot represent the down-regulated (light blue) and up-regulated (maroon) sQTLs with statistical significance in **A)** AD compared to controls, **B)** PD compared to controls, and **C)** AD compared to PD. The observed $-\log_{10}$ of P-values ($-\log_{10}(p)$) for *cis*-sQTLs related AS-events (Y-axis) are plotted against (Δ) PSI threshold (X-axis). Genes with $-\log_{10}(p) \leq 10$ are indicated.

An overlap analysis revealed 6,884 regions around *cis*-sQTLs variants in common between DLPDFC/BA9 and WB tissues from PD. The top enriched pathways were “establishment of protein localization to membrane” (enrichment ratio= 1.4985, $p \leq 8.3596\text{e-}10$) that included: *SYTL2*, *VAMP1*, and *VAMP3* (Table 4.5).

There were 817,540 *cis*-SNP-AS pairs associated with AS regulation or alteration of AS events for a specific gene in AD-WB samples compared to PD (Figure 4.10, Figure 4.11 C). The top enriched pathways were “positive regulation of cellular protein localization” (enrichment ratio= 1.7011, $p \leq 2.3667\text{e-}9$) that included *FYN*, and “cell cycle process” (enrichment ratio= 1.2681, $p \leq 2.2\text{e-}16$) that included *HLA-G* and “cellular response to stress” (enrichment ratio= 1.2652, $p \leq 2.2\text{e-}16$) that included *IL-37* (Table 4.4). No overlapping regions were found around the *cis*-sQTLs variants between AD and PD in WB tissue.

4.3.3 Comparative analysis of *cis*-eQTLs and *cis*-sQTLs identified via RNA-seq and genome sequencing from distinct cohorts

In this investigation, the RNA-seq was utilised to identify a set of QTLs, detailed comprehensively in Table 4.1. To verify the authenticity and reproducibility of our identified *cis*-eQTLs and *cis*-sQTLs, the results were benchmarked against genomic-based QTLs, summing up to approximately 3,785 entries, as reported in distinguished studies by Raj et al. (2018, Nature Genetics) and Li et al. (2019, Nature Communications).

A comparative breakdown of the overlap between our QTLs and the genomic-based ones is as below:

For AD Samples:

DLPFC/BA9: eQTLs (n= 115), sQTLs (n= 16)

PBMC: eQTLs (n= 6), sQTLs (n= 16)

WB: sQTLs (n= 8)

For PD Samples:

DLPFC/BA9: eQTLs (n= 80), sQTLs (n= 25)

PBMC: eQTLs (n= 36), sQTLs (n= 27)

WB: eQTLs (n= 86), sQTLs (n= 25)

Furthermore, it is pivotal to underscore that discrepancies in overlap between the studies could arise from a myriad of reasons. These could span the gamut from variations in the genetic makeup across populations, differences in the cohort sizes, the innate technical nuances pertaining to the respective sequencing methods, to even the overarching environmental factors influencing gene expression.

4.4 Discussions

This study analysed data sets of human central brain and peripheral blood transcriptome in combination with the genotyping data to identify eQTL, and sQTL SNPs associated with AD and PD in a genome-wide manner. The framework predicted thousands of functional SNPs for each GWAS locus. The presence of overlaps between the GWAS, and *cis*-eQTL/ -sQTL outputs means that some SNPs were related to NDs due to the change in AS events and of expression levels. In addition, many of the overlapped variants found to have sQTLs for a gene and eQTLs for a different one, while there were some variants involved simultaneously in the regulation of both

processes. This is not surprising, given that there is a substantial interplay between the molecular mechanisms underlying splicing and transcription, and because splicing often takes place co-transcriptionally [293].

In the AD-DLPFC, both *cis*-eQTLs and -sQTLs effects were identified mainly for cytoskeletal or membrane-associated genes (e.g., WASF1) [274], molecular machinery involved in clathrin-mediated endocytosis of the synaptic vesicles (e.g., the dynamin proteins (DNM1 & 2)) [275], and *mef2c*, a transcriptional regulatory factor, which mediates synaptic transmission and microglia activation in neurons [189]. As described in **Chapter 3**, molecular activity of all these genes were shown to be regulated by cyclin-dependent kinase 5 (Cdk5)); thus, making Cdk5 a highly possible candidate for development of therapeutics against AD.

For PD-DLPFC, both *cis*-eQTLs and -sQTLs effects were found for previously reported genes coding SNARE proteins and their interacting partners, such as STX1A, VAMP1, and SNAP25 [294]. ‘Interestingly’, the brain-derived diagnostic SNARE molecules were observed peripherally in PD. In this context, STX7, STX8, STX11, SYTL1, and VAPA were shown to have stronger associations for *cis*-sQTLs, while STX1A, STX16, STX17, STX18, STXBP4, STXBP5, and SYTL2 were among variants affected by gene expression. VAMP1, and VAMP2 were the only variant affected by both the splicing patterns, and expression. This outcome is in complete consistency with the previous result reported in **Chapter 2 & 3** that showed the role of SNARE complex formation defects in the pathogenesis of synucleinopathies in PD.

Little overlap here, the peripheral blood (PBMC and WB) and central brain in AD were similarly characterised with previously observed GWAS variants especially enriched among *cis*-sQTLs with roles in amyloidogenic pathway regulation and tau phosphorylation (e.g., Fyn) [191, 295], synapse formation and organization, synaptic cell adhesion, dendritic spine maturation and synaptic vesicle

release (e.g., Shank3) [296] and proinflammatory process mediated by microglia (e.g., IL-17RD) [297]. As described in **Chapter 3** altered expression of these genes was shown to be associated with mis-regulated activation of ERK 1/2 signalling cascade that led to A β aggregation and p-tau pathology in AD [248, 250, 295, 296].

In the brain of both NDs, *cis*-sQTLs were enriched with the previously reported GWAS variant ‘VEGF-A’ specifically associated with a broad array of signalling pathways contributing to the growth and maintenance of both vascular and neural cells (e.g., angiogenesis, neurogenesis, and neuroprotection) [298, 299], and blood-brain-barrier (BBB) permeability [237]. As reported in **Chapter 3** previous studies have linked mis-regulation of VEGF-A with BBB disruption shown to be associated with both AD and PD progression.

In the AD-PBMC, the *cis*-sQTLs were enriched in variants associated with inhibition of APP maturation and A β production and regulation of T cell activation (e.g., Homer3) [243, 300]. In addition, a previously observed GWAS variant ‘FC γ R2A’, an activator of phagocytosis in response to IgG-containing immune complexes [301], was especially enriched among the *cis*-eQTLs in AD blood (PBMC, and WB). Indeed, the *cis*-QTLs that differentiate AD from PD peripherally were co-localised with loci affecting both the splicing patterns and expression of genes related to major histocompatibility complex components (e.g., HLA-A, HLA-B, and HLA-DMB) [302]. Moreover, the previously GWAS observe immunity variants (HLA-DRB1, HLA-DPB1, and HLA-DQB1) [302] were especially enriched among *cis*-eQTLs in AD blood. As previously described in *Chapter 2 & 3* these findings support the idea that anomalies in peripheral immune response may function as a mediator in AD pathogenesis.

A final finding here, the peripheral PBMC in both NDs was characterized with the *cis*-sQTLs co-localised with loci affecting the splicing patterns of genes related to phosphoinositides-associated

proteins (e.g., INPP5B, and INPP5D). These proteins are known to play important roles in vesicular membrane trafficking, clathrin-mediated endocytosis and macroautophagy [303]. However, as reported in **Chapter 3** there are differences in PIPs molecular mechanism that contribute towards the development of neurological pathologies in both NDs.

In summary, the present results provide biological support to some of our previously identified GWAS variants associated with AD and PD in the peripheral blood, and central brain; thus, made it possible to create new hypotheses for increased risk of NDs via molecular effects. However, additional computational tests and experimental follow-ups are required for further confirmation of the current outputs.

4.5 Potential imitations and future prospective

Our study had several limitations. For instance, the comparisons of associated genes in the blood and brain may yield false results because they are based on independent groups ascertained from different studies, which were not matched for age, sex, ethnicity, and other factors that could consequently bias our findings, especially in the direction of false-negative QTLs mappings. A joint analysis across tissues, including multiple RNA-seq samples per individual would further improve the power for the future identification of QTLs. Secondly, very small sample sizes used in our study are likely to affect the power of QTLs identification in tissues and to induce false positive results in QTLs mapping. Therefore, usage of larger sample sizes would allow the detection of more tissue-specific QTLs with lower false discovery rates. Additionally, heterogeneous bulk RNA-seq samples used in the current research known to be comprised of diverse cell types that limits the power, interpretation, and downstream applications of QTL studies. Therefore, future work at the single-cell level will provide valuable insight into tissue/cell-type-associated variants and their downstream molecular effects that might be missed by

corresponding bulk RNA-seq studies. Furthermore, it was difficult to exclusively evaluate the roles of QTL variants in NDs, since multiple functional variants have effects on the same gene, which occasionally occurs in complex traits. In future work, a more detailed analysis of the relationship between QTLs and GWAS variants could be achieved by the usage of statistical methods to assess co-localization, and subsequent fine mapping of the QTL candidates to assign causal probabilities.

Chapter 5 General conclusions

In neurodegenerative disorders, the discovery of highly predictive genes and molecular pathways in both peripheral and brain tissues is imperative, as intervention needs to occur before the appearance of cognitive symptoms. To find genes that influence Alzheimer's disease (AD) and Parkinson's disease (PD) pathogenesis, human RNA-seq transcriptomic data from Brodmann Area 9 (BA9) of the dorsolateral prefrontal cortex (DLPFC), whole blood (WB), and peripheral blood mononuclear cells (PBMC) were analysed using a combination of differential gene expression and a random forest-based machine learning algorithm. The results suggest a slight overlap between PD and AD, and the AD brain signature is unique mainly compared to blood-based samples. The AD-DLPFC/BA9 was characterised by changes in *Mef2C*, and *TMED7* was shown to have roles in regulating microglia activation and TLR4 signalling. The peripheral AD transcriptome was associated with alterations in the 'viral process', and *FYN*, previously shown to link amyloid-beta and tau, was the prominent feature. However, in the absence of any overlap with the central transcriptome, it is unclear whether peripheral *FYN* levels reflect AD severity or progression. Additionally, AD-WB was flagged by changes in innate immune signalling and specifically TLR4/NF- κ B activation. Considering both the brain and blood biomarkers in AD, signatures support the hypothesis that changes in immune responses and microglia function is related to AD pathology. On the other hand, PD central and peripheral signatures are characterised by anomalies in 'exocytosis' and specific genes related to the SNARE complex, including *VAMP2*, *STX1A*, and *PAK1*. This is consistent with our current understanding of the physiological role of alpha-synuclein and how alpha-synuclein oligomers compromise vesicle docking and neurotransmission. Overall, the results describe distinct disease-specific pathomechanisms within the brain and peripherally for the two most common neurodegenerative disorders.

As for many genetically heterogeneous and complex diseases, the application of genome-wide association screening to identify susceptibility genes has gained considerable momentum; the follow-up study was designed to compare the specified results with GWAS outputs. Hence, RNA-seq data from the same samples were used for SNP calling to identify genetic variants important for describing different risk phenotypes in AD and PD.

The results suggested that dysregulation of brain angiogenesis biomarker (e.g., VEGF-A) and blood-related phosphoinositides-associated proteins might lead to abnormalities in AD as well as PD. Previous studies have linked mis-regulation of VEGF-A with BBB disruption shown to be associated with both AD and PD progression. However, there are differences in PIPs molecular mechanism that contribute towards the development of neurological pathologies in both NDs.

Not directly here, deregulation of Cdk5 in the brain might contribute to an array of pathological events in AD, since the function of GWAS-associated risk genes discovered in AD-brain (e.g., *mef2c*, *DNM2*, and *WASF1*) were shown to be regulated by Cdk5. This finding suggests an exciting area of Cdk5 biology remain to be explored in AD.

The genome-wide approach's exciting finding repeated the previous results for PD-DLPFC/BA9 by revealing multiple genes associated with v- and t-SNAREs (*VAMP2*, *STX1A*, and *SYT2*) and Qabc-SNAREs such as *VAMP3*, *STX7*, *STX8*, and *SYT7*. Simultaneously, these genes were identified in PD cases peripherally. This evidence supports a role for the relationship between α -syn and SNARE protein redistribution and synaptic failure in the PD process.

In this work I identified new common signatures between the AD brain and the blood (e.g., *SHANK3*, *Fyn*, and family of IL-17 receptors), which have been shown to contribute to amyloid-beta and tau pathology by activating ERK 1/2 signalling cascade. This finding suggested that the

ERK 1/2 pathway might be a significant tool and a valid therapeutic target in the future studies related to AD pathology.

Consistent with the previous results, AD blood was characterized with immune response factors, including Homer2 & 3, FC γ R2A, and HLA class I, and II that may support a role for anomalies in peripheral immune response and their function as a mediator in AD pathogenesis.

Overall, while repeating some of the previous results, GWAS analysis revealed proportions of overlaps between AD and PD in both the central brain and peripheral blood. An additional finding was the discovery of potential blood-brain AD biomarkers that highlight the contribution of different genes in the aggregation of amyloid- β and p-tau, which are significantly associated with AD risk.

To found additional evidence for the role of GWAS-associated markers in AD and PD pathology, QTL analysis was performed using RNA-seq alternative splicing, and gene expression data and SNPs discovered from the same RNA-seq samples. Comparison of the significant *cis*-eQTLs/ -sQTLs with the GWAS-associated SNPs identified overlapped loci, and genes that were shown previously to be linked to AD and PD pathology. The presence of overlaps between GWAS, *cis*-eQTL, and *cis*-sQTL, implicate changes in AS events and expression patterns of ND-related risk genes that are likely responsible for modulating disease susceptibility in central brain and peripheral blood of AD and PD. In addition, many overlapped variants were found to have sQTLs for a gene and eQTLs for a different one. At the same time, there were some variants involved simultaneously in the regulation of both processes. This is not surprising, given that there is a substantial interplay between the molecular mechanisms underlying splicing and transcription, and because splicing often takes place co-transcriptionally [276].

From a transcriptional perspective, our findings underscore the necessity of subsequent investigations geared toward the accurate identification and validation of therapeutic targets, ones that have previously delineated within the context of disease-specific pathomechanisms for AD and PD - two of the most pervasive neurodegenerative disorders.

Notably, incorporating blood-derived datasets in our research provides a unique lens through which peripheral biomarkers can be ascertained, offering invaluable insights into systemic responses and potential diagnostic or therapeutic avenues. Complemented by thoroughly examining overlaps, this approach ensures a comprehensive understanding of shared and distinct molecular underpinnings across these disorders.

Incorporating data from GWAS and QTL analyses, this scholarly endeavour elucidates intricate relationships underlying complex neurodegenerative diseases, notably AD and PD. The interconnectedness and convergence of these data sources offer unprecedented depth and granularity in our understanding.

As we gaze into the horizon of neurodegenerative research, the integration of advanced technological platforms holds paramount significance. Future inquiries, fortified by the amalgamation of long-read sequencing, single-cell RNA sequencing, proteomic analysis, and state-of-the-art neuroimaging modalities, offer a promising avenue. Such integrative approaches, buttressed by computational stringency, promise enhanced detection and quantification capabilities, particularly in discerning cell-type-specific neurodegenerative risk alleles in both circulatory and neural matrices.

5.1 Differentiation from previous works on gene associations with neurodegenerative disorders (NDs).

In the quest to unravel NDs' genetic and molecular intricacies, numerous researchers have embarked on comprehensive studies to unveil associations between genes and disease

pathogenesis. A notable endeavour in this domain emerges from *Philip L. De Jager* lab, which has persistently sought to discern pivotal genetic associations. However, NDs' vast and convoluted landscape mandates continual exploration, leveraging innovative methodologies and interdisciplinary approaches.

This study presents a multifaceted approach that melds RNA-seq transcriptomic data analyses with machine learning algorithms to uncover potential gene interactions in AD and PD. While the use of RNA-seq is not novel, the combination with machine learning algorithms, differential gene expression, and the subsequent linkage to genome-wide association screening delve deeper into the intricacies of the disorders than most previous studies.

Several salient features and findings distinguish the current research:

Peripheral and Central Pathomechanisms: A holistic understanding of disease-specific pathomechanisms have been provided centrally and peripherally by analysing transcriptomic data from diverse regions and tissues.

Overlaps and Distinctions: The results illuminated overlaps between PD and AD but, more crucially, highlighted the unique PD brain signature.

Incorporation of GWAS Data: The innovative use of GWAS data alongside RNA-seq has unveiled intriguing overlaps that bear significant implications for disease understanding and potential therapeutic targets.

Rigorous Validation through Genomic Comparisons: By juxtaposing the RNA-seq-derived QTLs with genomic-based QTLs from renowned studies, I sought to ensure the authenticity and reproducibility of our findings.

Emphasis on Interactions and Mechanisms: The findings stress the interplay between splicing and transcription, underlining the significance of understanding these processes for NDs.

In sum, while my research stands on the shoulders of previous giants in the field, it undeniably introduces a fresh perspective, robust methodology, and valuable insights into the intricate world of AD and PD. As NDs continue challenging global health, such integrative approaches are desirable and indispensable.

References

1. McAlary L, Plotkin SS, Cashman NR. Emerging developments in targeting proteotoxicity in neurodegenerative diseases. *CNS Drugs*. 2019;33(9).883-904.
2. Erkkinen, M.G., Kim, M.O. and Geschwind, M.D. Clinical neurology, and epidemiology of the major neurodegenerative diseases. *Cold Spring Harbor perspectives in biology*. 2018;10(4).033118.
3. M Ashraf G, H Greig N, A Khan T, Hassan I, Tabrez S, Shakil S, A Sheikh I, K Zaidi S, Akram M, R Jabir N, K Firoz C. Protein misfolding and aggregation in Alzheimer's disease and type 2 diabetes mellitus. *CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)*. 2014;13(7).1280-1293.
4. Mannini B, Chiti F. Chaperones as suppressors of protein misfolded oligomer toxicity. *Frontiers in molecular neuroscience*. 2017;10.98.
5. Gomez-Gutierrez R, Morales R. The prion-like phenomenon in Alzheimer's disease: Evidence of pathology transmission in humans. *Advances in Medical Biochemistry, Genomics, Physiology, and Pathology*. 2021.639-647.
6. José A, Ferrer I, Gavín R. Role of cellular prion protein in interneuronal amyloid transmission. *Progress in Neurobiology*. 2018;165.87-102.
7. Kovacs GG. Molecular pathology of neurodegenerative diseases: principles and practice. *Journal of clinical pathology*. 2019;72(11).725-35.
8. Agrawal M, Biswas A. Molecular diagnostics of neurodegenerative disorders. *Frontiers in molecular biosciences*. 2015;2.54.
9. Courtney E, Kornfeld S, Janitz K, Janitz M. Transcriptome profiling in neurodegenerative disease. *Journal of neuroscience methods*. 2010;193(2).189-202.
10. Sun H, Shen XR, Fang ZB, Jiang ZZ, Wei XJ, Wang ZY, Yu XF. Next-generation sequencing technologies and neurogenetic diseases. *Life*. 2021;11(4).361.
11. Hrdlickova R, Toloue M, Tian B. RNA-Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews: RNA*. 2017;8(1).e1364.
12. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nature Reviews Genetics*. 2019;20(11).631-656.

13. Dündar F, Skrabanek L, Zumbo P. Introduction to differential gene expression analysis using RNA-seq. *Applied Bioinformatics Core/Weill Cornell Medical College*. 2015;1-67.
14. McDermaid A, Monier B, Zhao J, Liu B, Ma Q. Interpretation of differential gene expression results of RNA-seq data: review and integration. *Briefings in bioinformatics*. 2019;20(6).2044-2054.
15. Abbas M, El-Manzalawy Y. Machine learning based refined differential gene expression analysis of pediatric sepsis. *BMC medical genomics*. 2020;13(1).1-10.
16. Ahsan MM, Luna SA, Siddique Z. Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. *InHealthcare*. 2022;10(3).541.
17. Myszczyńska MA, Ojamies PN, Lacoste A, Neil D, Saffari A, Mead R, Hautbergue GM, Holbrook JD, Ferraiuolo L. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nature Reviews Neurology*. 2020;16(8).440-456.
18. Hira ZM, Gillies DF. A review of feature selection and feature extraction methods applied on microarray data. *Advances in bioinformatics*. 2015.
19. Su K, Yu T, Wu H. Accurate feature selection improves single-cell RNA-seq cell clustering. *Briefings in bioinformatics*. 2021;22(5).bbab034.
20. Mills JD, Janitz M. Alternative splicing of mRNA in the molecular pathology of neurodegenerative diseases. *Neurobiology of aging*. 2012;33(5).1012-e11.
21. Jiang W, Chen L. Alternative splicing: Human disease and quantitative analysis from high-throughput sequencing. *Computational and Structural Biotechnology Journal*. 2021;19.183-195.
22. Love JE, Hayden EJ, Rohn TT. Alternative splicing in Alzheimer's disease. *Journal of Parkinson's disease and Alzheimer's disease*. 2015;2(2).
23. Halperin RF, Hegde A, Lang JD, Raupach EA, Legendre C, Liang WS, LoRusso PM, Sekulic A, Sosman JA, Trent JM, Rangasamy S. Improved methods for RNAseq-based alternative splicing analysis. *Scientific reports*. 2021;11(1).1-15.
24. Zhao Y, Wang K, Wang WL, Yin TT, Dong WQ, Xu CJ. A high-throughput SNP discovery strategy for RNA-seq data. *BMC genomics*. 2019;20(1).1-10.
25. Robert F, Pelletier J. Exploring the impact of single-nucleotide polymorphisms on translation. *Frontiers in genetics*. 2018;9.507.

26. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*. 2019;20(8).467-484.
27. Jehl F, Degalez F, Bernard M, Lecerf F, Lagoutte L, Désert C, Coulée M, Bouchez O, Leroux S, Abasht B, Tixier-Boichard M. RNA-seq data for reliable SNP detection and genotype calling: interest for coding variant characterization and cis-regulation analysis by allele-specific expression in livestock species. *Frontiers in genetics*. 2021;12.1104.
28. Bakhtiarizadeh MR, Alamouti AA. RNA-seq based genetic variant discovery provides new insights into controlling fat deposition in the tail of sheep. *Scientific Reports*. 2020;10(1).1-13.
29. Hertrich I, Dietrich S, Blum C, Ackermann H. The role of the dorsolateral prefrontal cortex for speech and language processing. *Frontiers in Human Neuroscience*. 2021; 17(15).217.
30. Kumar S, Zomorodi R, Ghazala Z, Goodman MS, Blumberger DM, Cheam A, Fischer C, Daskalakis ZJ, Mulsant BH, Pollock BG, Rajji TK. Extent of dorsolateral prefrontal cortex plasticity and its association with working memory in patients with Alzheimer disease. *JAMA psychiatry*. 2017;74(12).1266-74.
31. Caspers J, Mathys C, Hoffstaedter F, Südmeyer M, Cieslik EC, Rubbert C, Hartmann CJ, Eickhoff CR, Reetz K, Grefkes C, Michely J. Differential functional connectivity alterations of two subdivisions within the right dlPFC in Parkinson's disease. *Frontiers in human neuroscience*. 2017;11.288.
32. Augustine J, Jereesh AS. Blood-based gene-expression biomarkers identification for the non-invasive diagnosis of Parkinson's disease using two-layer hybrid feature selection. *Gene*. 2022;823.146366.
33. Al-Bachari S, Naish JH, Parker GJ, Emsley HC, Parkes LM. Blood–brain barrier leakage is increased in Parkinson's disease. *Frontiers in physiology*. 2020;11.593026.
34. Sweeney MD, Sagare AP, Zlokovic BV. Blood–brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nature Reviews Neurology*. 2018;14(3).133-150.
35. Wang C, Chen L, Yang Y, Zhang M, Wong G. Identification of potential blood biomarkers for Parkinson's disease by gene expression and DNA methylation data integration analysis. *Clinical epigenetics*. 2019;11(1).1-15.
36. Ashton NJ, Hye A, Rajkumar AP, Leuzy A, Snowden S, Suárez-Calvet M, Karikari TK, Schöll M, La Joie R, Rabinovici GD, Höglund K. An update on blood-based biomarkers for non-Alzheimer neurodegenerative disorders. *Nature Reviews Neurology*. 2020;16(5).265-84.

37. Palmqvist S, Tideman P, Cullen N, Zetterberg H, Blennow K, Dage JL, Stomrud E, Janelidze S, Mattsson-Carlgren N, Hansson O. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nature Medicine*. 2021;27(6).1034-1042.
38. Foulds PG, Diggle P, Mitchell JD, Parker A, Hasegawa M, Masuda-Suzukake M, Mann D, Allsop D. A longitudinal study on α -synuclein in blood plasma as a biomarker for Parkinson's disease. *Scientific reports*. 2013;3(1).1-6.
39. Arosio B, D'Addario C, Gussago C, Casati M, Tedone E, Ferri E, Nicolini P, Rossi PD, Maccarrone M, Mari D. Peripheral blood mononuclear cells as a laboratory to study dementia in the elderly. *BioMed research international*. 2014.
40. Pansarasa O, Garofalo M, Scarian E, Dragoni F, Garau J, Di Gerlando R, Diamanti L, Bordoni M, Gagliardi S. Biomarkers in human peripheral blood mononuclear cells: The State of the art in amyotrophic lateral sclerosis. *International journal of molecular sciences*. 2022;23(5).2580.
41. Arosio B, D'Addario C, Gussago C, Casati M, Tedone E, Ferri E, Nicolini P, Rossi PD, Maccarrone M, Mari D. Peripheral blood mononuclear cells as a laboratory to study dementia in the elderly. *BioMed research international*. 2014;2014.
42. Kobo H, Bar-Shira A, Dahary D, Gan-Or Z, Mirelman A, Goldstein O, Giladi N, Orr-Urtreger A. Down-regulation of B cell-related genes in peripheral blood leukocytes of Parkinson's disease patients with and without GBA mutations. *Molecular Genetics and Metabolism*. 2016;117(2).179-85.
43. Jalilianhasanpour, R., Beheshtian, E., Sherbaf, G., Sahraian, S. and Sair, H.I. Functional connectivity in neurodegenerative disorders: Alzheimer's disease and frontotemporal dementia. *Topics in Magnetic Resonance Imaging*. 2019;28(6).317-324.
44. Milind N, Preuss C, Haber A, Ananda G, Mukherjee S, John C, Shapley S, Logsdon BA, Crane PK, Carter GW. Transcriptomic stratification of late-onset Alzheimer's cases reveals novel genetic modifiers of disease pathology. *PLoS Genetics*. 2020;16(6).1008775.
45. Simons JS, Spiers HJ. Prefrontal and medial temporal lobe interactions in long-term memory. *Nature reviews neuroscience*. 2003 Aug;4(8):637-48.
46. Rabinovici GD. Late-onset Alzheimer Disease. *Continuum (Minneap Minn)*. 2019;25(1).14-33.

47. Dong HK, Gim JA, Yeo SH, Kim HS. Integrated late onset Alzheimer's disease (LOAD) susceptibility genes: cholesterol metabolism and trafficking perspectives. *Gene*. 2017;597:10-16.
48. Yamazaki Y, Zhao N, Caulfield TR, Liu CC, Bu G. Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. *Nature Reviews Neurology*. 2019;15(9):501-18.
49. Ransohoff RM. How neuroinflammation contributes to neurodegeneration. *Science*. 2016;353(6301):777-783.
50. Guimas Almeida C, Sadat Mirfakhar F, Perdigão C, Burrinha T. Impact of late-onset Alzheimer's genetic risk factors on beta-amyloid endocytic production. *Cellular and Molecular Life Sciences*. 2018;75(14):2577-2589.
51. Reitz C, Rogaeva E, Beecham GW. Late-onset vs nonmendelian early-onset Alzheimer disease: A distinction without a difference? *Neurology Genetics*. 2020;6(5).
52. Sirkis DW, Bonham LW, Johnson TP, La Joie R, Yokoyama JS. Dissecting the clinical heterogeneity of early-onset Alzheimer's disease. *Molecular Psychiatry*. 2022;27(6):2674-2688.
53. Van Bulck M, Sierra-Magro A, Alarcon-Gil J, Perez-Castillo A, Morales-Garcia JA. Novel approaches for the treatment of Alzheimer's and Parkinson's disease. *International Journal of Molecular Sciences*.
54. Ayodele T, Rogaeva E, Kurup JT, Beecham G, Reitz C. Early-onset Alzheimer's disease: What is missing in research? *Current neurology and neuroscience reports*. 2021;21(2):1-10.
55. Coronel R, Bernabeu-Zornoza A, Palmer C, Muñoz-Moreno M, Zambrano A, Cano E, Liste I. Role of amyloid precursor protein (APP) and its derivatives in the biology and cell fate specification of neural stem cells. *Molecular neurobiology*. 2018;55(9):7107-7117.
56. Guo T, Zhang D, Zeng Y, Huang TY, Xu H, Zhao Y. Molecular and cellular mechanisms underlying the pathogenesis of Alzheimer's disease. *Molecular neurodegeneration*. 2020;15(1):1-37.
57. Coburger I, Dahms SO, Roeser D, Gührs KH, Hortschansky P, Than ME. Analysis of the overall structure of the multi-domain amyloid precursor protein (APP). *PLoS One*. 2013;8(12):81926.
58. Mañucat-Tan NB, Saadipour K, Wang YJ, Bobrovskaya L, Zhou XF. Cellular trafficking of amyloid precursor protein in Amyloidogenesis physiological and pathological significance. *Molecular Neurobiology*. 2019;56(2):812-830.

59. Jiang S, Li Y, Zhang X, Bu G, Xu H, Zhang YW. Trafficking regulation of proteins in Alzheimer's disease. *Molecular neurodegeneration*. 2014;9(1).1-13.
60. Zhang X, Song W. The role of APP and BACE1 trafficking in APP processing and amyloid- β generation. *Alzheimer's research & therapy*. 2013;5(5).1-8.
61. Lichtenthaler SF. Alpha-secretase in Alzheimer's disease: molecular identity, regulation, and therapeutic potential. *Journal of neurochemistry*. 2011;116(1).10-21.
62. Manzine PR, Ettcheto M, Cano A, Busquets O, Marcello E, Pelucchi S, Di Luca M, Endres K, Olloquequi J, Camins A, Cominetti MR. ADAM10 in Alzheimer's disease: Pharmacological modulation by natural compounds and its role as a peripheral marker. *Biomedicine & Pharmacotherapy*. 2019;113.108661.
63. Wang X, Zhou X, Li G, Zhang Y, Wu Y, Song W. Modifications, and trafficking of APP in the pathogenesis of Alzheimer's disease. *Frontiers in molecular neuroscience*. 2017;10.294.
64. Tan J, Evin G. β -Site APP-cleaving enzyme 1 trafficking and Alzheimer's disease pathogenesis. *Journal of neurochemistry*. 2012;120(6).869-880.
65. Matz A, Halamoda-Kenzaoui B, Hamelin R, Mosser S, Alattia JR, Dimitrov M, Moniatte M, Fraering PC. Identification of new Presenilin-1 phosphosites: implication for γ -secretase activity and A β production. *Journal of Neurochemistry*. 2015;133(3).409-421.
66. Dries DR, Yu G. Assembly, maturation, and trafficking of the γ -secretase complex in Alzheimer's disease. *Current Alzheimer Research*. 2008;5(2).132-146.
67. Coronel R, Bernabeu-Zornoza A, Palmer C, Muñiz-Moreno M, Zambrano A, Cano E, Liste I. Role of amyloid precursor protein (APP) and its derivatives in the biology and cell fate specification of neural stem cells. *Molecular neurobiology*. 2018;55(9).7107-7117.
68. Tan JZ, Gleeson PA. The role of membrane trafficking in the processing of amyloid precursor protein and production of amyloid peptides in Alzheimer's disease. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. 2019;1861(4).697-712.
69. Müller UC, Deller T, Korte M. Not just amyloid: physiological functions of the amyloid precursor protein family. *Nature Reviews Neuroscience*. 2017;18(5).281-298.
70. Walsh DM, Selkoe DJ. Amyloid β -protein and beyond: the path forward in Alzheimer's disease. *Current opinion in neurobiology*. 2020;61.116-124.

71. Multhaup G, Huber O, Buée L, Galas MC. Amyloid precursor protein (APP) metabolites APP intracellular fragment (AICD), A β 42, and tau in nuclear roles. *Journal of Biological Chemistry*. 2015;290(39).23515-23522.
72. Chen GF, Xu TH, Yan Y, Zhou YR, Jiang Y, Melcher K, Xu HE. Amyloid beta: structure, biology, and structure-based therapeutic development. *Acta Pharmacologica Sinica*. 2017;38(9).1205-1235.
73. de Coninck D, Schmidt TH, Schloetel JG, Lang T. Packing density of the amyloid precursor protein in the cell membrane. *Biophysical journal*. 2018;114(5).1128-41.
74. Hampel H, Hardy J, Blennow K, Chen C, Perry G, Kim SH, Villemagne VL, Aisen P, Vendruscolo M, Iwatsubo T, Masters CL. The amyloid- β pathway in Alzheimer's disease. *Molecular Psychiatry*. 2021;26(10).5481-5503.
75. Spitzer P, Condic M, Herrmann M, Oberstein TJ, Scharin-Mehlmann M, Gilbert DF, Friedrich O, Grömer T, Kornhuber J, Lang R, Maler JM. Amyloidogenic amyloid- β -peptide variants induce microbial agglutination and exert antimicrobial activity. *Scientific reports*. 2016;6(1).1-11.
76. Kumar DK, Choi SH, Washicosky KJ, Eimer WA, Tucker S, Ghofrani J, Lefkowitz A, McColl G, Goldstein LE, Tanzi RE, Moir RD. Amyloid- β peptide protects against microbial infection in mouse and worm models of Alzheimer's disease. *Science translational medicine*. 2016;8(340).340ra72.
77. Torrent M, Pulido D, Nogues MV, Boix E. Exploring new biological functions of amyloids: bacteria cell agglutination mediated by host protein aggregation. *PLoS pathogens*. 2012;8(11).1003005.
78. Alonso R, Pisa D, Fernández-Fernández AM, Carrasco L. Infection of fungi and bacteria in brain tissue from elderly persons and patients with Alzheimer's disease. *Frontiers in aging neuroscience*. 2018;10.159.
79. Moir RD, Lathe R, Tanzi RE. The antimicrobial protection hypothesis of Alzheimer's disease. *Alzheimer's & Dementia*. 2018;14(12).1602-1614.
80. Gallego Villarejo L, Bachmann L, Marks D, Brachthäuser M, Geidies A, Müller T. Role of Intracellular Amyloid β as Pathway Modulator, Biomarker, and Therapy Target. *International Journal of Molecular Sciences*. 2022;23(9).4656.
81. Kaye R, Lasagna-Reeves CA. Molecular mechanisms of amyloid oligomers toxicity. *Journal of Alzheimer's Disease*. 2013;33(s1).67-78.

82. Mroczko B, Groblewska M, Litman-Zawadzka A, Kornhuber J, Lewczuk P. Cellular receptors of amyloid β oligomers (A β Os) in Alzheimer's disease. *International journal of molecular sciences*. 2018;19(7).1884.
83. Huang YR, Liu RT. The toxicity and polymorphism of β -amyloid oligomers. *International journal of molecular sciences*. 2020;21(12).4477.
84. Sun X, Chen WD, Wang YD. β -Amyloid: the key peptide in the pathogenesis of Alzheimer's disease. *Frontiers in pharmacology*. 2015;6.221.
85. Ganguly G, Chakrabarti S, Chatterjee U, Saso L. Proteinopathy, oxidative stress and mitochondrial dysfunction: cross talk in Alzheimer's disease and Parkinson's disease. *Drug design, development, and therapy*. 2017;11.797.
86. Karikari TK, Ashton NJ, Brinkmalm G, Brum WS, Benedet AL, Montoliu-Gaya L, Lantero-Rodriguez J, Pascoal TA, Suárez-Calvet M, Rosa-Neto P, Blennow K. Blood phospho-tau in Alzheimer disease: analysis, interpretation, and clinical utility. *Nature Reviews Neurology*. 2022.1-19.
87. D 'Errico P, Meyer-Luehmann M. Mechanisms of pathogenic tau and A β protein spreading in Alzheimer's disease. *Frontiers in Aging Neuroscience*. 2020;12.265.
88. Sinsky J, Pichlerova K, Hanes J. Tau protein interaction partners and their roles in Alzheimer's disease and other tauopathies. *International Journal of Molecular Sciences*. 2021;22(17).9207.
89. Kent SA, Spires-Jones TL, Durrant CS. The physiological roles of tau and A β : implications for Alzheimer's disease pathology and therapeutics. *Acta neuropathologica*. 2020;140(4).417-447.
90. Vaz-Silva J, Gomes P, Jin Q, Zhu M, Zhuravleva V, Quintremil S, Meira T, Silva J, Dioli C, Soares-Cunha C, Daskalakis NP. Endolysosomal degradation of Tau and its role in glucocorticoid-driven hippocampal malfunction. *The EMBO journal*. 2018;37(20).e99084.
91. Jiang S, Bhaskar K. Degradation and transmission of tau by autophagic-endolysosomal networks and potential therapeutic targets for tauopathy. *Frontiers in Molecular Neuroscience*. 2020;13.586731.
92. Busche MA, Hyman BT. Synergy between amyloid- β and tau in Alzheimer's disease. *Nature neuroscience*. 2020;23(10).1183-1193.
93. Pickett EK, Herrmann AG, McQueen J, Abt K, Dando O, Tulloch J, Jain P, Dunnett S, Sohrabi S, Fjeldstad MP, Calkin W. Amyloid beta and tau cooperate to cause reversible behavioral and transcriptional deficits in a model of Alzheimer's disease. *Cell reports*. 2019;29(11).3592-3604.

94. Irvine GB, El-Agnaf OM, Shankar GM, Walsh DM. Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. *Molecular medicine*. 2008;14(7).451-464.
95. Rajmohan R, Reddy PH. Amyloid-beta and phosphorylated tau accumulations cause abnormalities at synapses of Alzheimer's disease neurons. *Journal of Alzheimer's Disease*. 2017;57(4).975-99.
96. Yang YH, Huang LC, Hsieh SW, Huang LJ. Dynamic blood concentrations of A β 1–40 and A β 1–42 in Alzheimer's disease. *Frontiers in Cell and Developmental Biology*. 2020;8.768.
97. Ramanathan A, Nelson AR, Sagare AP, Zlokovic BV. Impaired vascular-mediated clearance of brain amyloid beta in Alzheimer's disease: the role, regulation, and restoration of LRP1. *Frontiers in aging neuroscience*. 2015;7.136.
98. Guimas Almeida C, Sadat Mirfakhar F, Perdigão C, Burrinha T. Impact of late-onset Alzheimer's genetic risk factors on beta-amyloid endocytic production. *Cellular and Molecular Life Sciences*. 2018;75(14).2577-2589.
99. Cuchillo-Ibañez I, Lennol MP, Escamilla S, Mata-Balaguer T, Valverde-Vozmediano L, Lopez-Font I, Ferrer I, Sáez-Valero J. The apolipoprotein receptor LRP3 compromises APP levels. *Alzheimer's Research & Therapy*. 2021;13(1).1-17.
100. Storck SE, Hartz AM, Bernard J, Wolf A, Kachlmeier A, Mahringer A, Weggen S, Pahnke J, Pietrzik CU. The concerted amyloid-beta clearance of LRP1 and ABCB1/P-gp across the blood-brain barrier is linked by PICALM. *Brain, behavior, and immunity*. 2018;73.21-33.
101. Zhang YL, Wang J, Zhang ZN, Su Q, Guo JH. The relationship between amyloid-beta and brain capillary endothelial cells in Alzheimer's disease. *Neural Regeneration Research*. 2022;17(11).2355.
102. Zuroff L, Daley D, Black KL, Koronyo-Hamaoui M. Clearance of cerebral A β in Alzheimer's disease: reassessing the role of microglia and monocytes. *Cellular and Molecular Life Sciences*. 2017;74(12).2167-2201.
103. Menzies FM, Fleming A, Caricasole A, Bento CF, Andrews SP, Ashkenazi A, Füllgrabe J, Jackson A, Sanchez MJ, Karabiyik C, Licitra F. Autophagy, and neurodegeneration: pathogenic mechanisms and therapeutic opportunities. *Neuron*. 2017;93(5).1015-1034.

104. Long Z, Chen J, Zhao Y, Zhou W, Yao Q, Wang Y, He G. Dynamic changes of autophagic flux induced by Abeta in the brain of postmortem Alzheimer's disease patients, animal models and cell models. *Aging* (Albany NY). 2020;12(11).10912.
105. Cuddy LK, Alia AO, Salvo MA, Chandra S, Grammatopoulos TN, Justman CJ, Lansbury PT, Mazzulli JR, Vassar R. Farnesyltransferase inhibitor LNK-754 attenuates axonal dystrophy and reduces amyloid pathology in mice. *Molecular neurodegeneration*. 2022;17(1).1-23.
106. Sharma J, Di Ronza A, Lotfi P, Sardiello M. Lysosomes and brain health. *Annual review of neuroscience*. 2018;41.255-276.
107. Domingues HS, Portugal CC, Socodato R, Relvas JB. Oligodendrocyte, astrocyte, and microglia crosstalk in myelin development, damage, and repair. *Frontiers in cell and developmental biology*. 2016;4.71.
108. Leng F, Edison P. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here? *Nature Reviews Neurology*. 2021;17(3).157-172.
109. Colonna M, Butovsky O. Microglia function in the central nervous system during health and neurodegeneration. *Annual review of immunology*. 2017;35.441.
110. Reed-Geaghan EG, Savage JC, Hise AG, Landreth GE. CD14 and toll-like receptors 2 and 4 are required for fibrillar A β -stimulated microglial activation. *Journal of Neuroscience*. 2009;29(38).11982-11992.
111. Heneka MT, Golenbock DT, Latz E. Innate immunity in Alzheimer's disease. *Nature immunology*. 2015;16(3).229-236.
112. Milner MT, Maddugoda M, Götz J, Burgener SS, Schroder K. The NLRP3 inflammasome triggers sterile neuroinflammation and Alzheimer's disease. *Current Opinion in Immunology*. 2021;68.116-124.
113. Heneka MT, McManus RM, Latz E. Inflammasome signalling in brain function and neurodegenerative disease. *Nature Reviews Neuroscience*. 2018;19(10).610-621.
114. Ghosh S, Wu MD, Shafteel SS, Kyrkanides S, LaFerla FM, Olschowka JA, O'Banion MK. Sustained interleukin-1 β overexpression exacerbates tau pathology despite reduced amyloid burden in an Alzheimer's mouse model. *Journal of Neuroscience*. 2013;33(11).5053-64.
115. Mehta NH, Suss RA, Dyke JP, Theise ND, Chiang GC, Strauss S, Saint-Louis L, Li Y, Pahlajani S, Babaria V, Glodzik L. Quantifying cerebrospinal fluid dynamics: A review of human neuroimaging contributions to CSF physiology and neurodegenerative disease. *Neurobiology of Disease*. 2022.105776.

- 116.Rasmussen MK, Mestre H, Nedergaard M. Fluid transport in the brain. *Physiological Reviews*. 2022;102(2).1025-1151.
- 117.Xia M, Yang L, Sun G, Qi S, Li B. Mechanism of depression as a risk factor in the development of Alzheimer's disease: the function of AQP4 and the glymphatic system. *Psychopharmacology*. 2017;234(3).365-379.
- 118.Christensen J, Li C, Mychasiuk R. Choroid plexus function in neurological homeostasis and disorders: The awakening of the circadian clocks and orexins. *Journal of Cerebral Blood Flow & Metabolism*. 2022.0271678X221082786.
- 119.Silva I, Silva J, Ferreira R, Trigo D. Glymphatic system, AQP4, and their implications in Alzheimer's disease. *Neurological Research and Practice*. 2021;3(1).1-9.
- 120.Harrison IF, Ismail O, Machhada A, Colgan N, Ohene Y, Nahavandi P, Ahmed Z, Fisher A, Meftah S, Murray TK, Ottersen OP. Impaired glymphatic function and clearance of tau in an Alzheimer's disease model. *Brain*. 2020;143(8).2576-2593.
- 121.Han, M., Nagele, E., DeMarshall, C., Acharya, N. and Nagele, R. Diagnosis of Parkinson's disease based on disease-specific autoantibody profiles in human sera. *PloS one*. 2012;7(2).p.e32383.
- 122.Tan JM, Wong ES, Lim KL. Protein misfolding and aggregation in Parkinson's disease. *Antioxidants & redox signalling*. 2009;11(9).2119-2134.
- 123.Salahuddin P, Fatima MT, Uversky VN, Khan RH, Islam Z, Furkan M. The role of amyloids in Alzheimer's and Parkinson's diseases. *International Journal of Biological Macromolecules*. 2021;190.44-55.
- 124.Padilla-Godínez FJ, Ramos-Acevedo R, Martínez-Becerril HA, Bernal-Conde LD, Garrido-Figueroa JF, Hiriart M, Hernández-López A, Argüero-Sánchez R, Callea F, Guerra-Crespo M. Protein misfolding and aggregation: The relatedness between Parkinson's disease and hepatic endoplasmic reticulum storage disorders. *International Journal of Molecular Sciences*. 2021;22(22).12467.
- 125.Teixeira M, Sheta R, Idi W, Oueslati A. Alpha-Synuclein and the endolysosomal system in parkinson's disease: guilty by association. *Biomolecules*. 2021;11(9).1333.
- 126.Chandra S, Chen X, Rizo J, Jahn R, Südhof TC. A broken α -helix in folded α -synuclein. *Journal of Biological Chemistry*. 2003;278(17).15313-15318.

127. Faustini G, Bono F, Valerio A, Pizzi M, Spano P, Bellucci A. Mitochondria and α -synuclein: friends or foes in the pathogenesis of Parkinson's disease? *Genes*. 2017;8(12).377.
128. Majd S, Power JH, Grantham HJ. Neuronal response in Alzheimer's and Parkinson's disease: the effect of toxic proteins on intracellular pathways. *BMC neuroscience*. 2015;16(1).1-13.
129. Park SC, Moon JC, Shin SY, Son H, Jung YJ, Kim NH, Kim YM, Jang MK, Lee JR. Functional characterization of alpha-synuclein protein with antimicrobial activity. *Biochemical and biophysical research communications*. 2016;478(2).924-928.
130. Chakraborty R, Dey S, Sil P, Paul SS, Bhattacharyya D, Bhunia A, Sengupta J, Chattopadhyay K. Conformational distortion in a fibril-forming oligomer arrests alpha-Synuclein fibrillation and minimizes its toxic effects. *Communications biology*. 2021;4(1).1-14.
131. Vidović M, Rikalovic MG. Alpha-synuclein aggregation pathway in Parkinson's disease: Current status and novel therapeutic approaches. *Cells*. 2022;11(11).1732.
132. Zhang J, Li X, Li JD. The Roles of post-translational modifications on α -synuclein in the pathogenesis of Parkinson's diseases. *Frontiers in neuroscience*. 2019;13.381.
133. Tyson T, Steiner JA, Brundin P. Sorting out release, uptake, and processing of alpha-synuclein during prion-like spread of pathology. *Journal of neurochemistry*. 2016;139.275-289.
134. Ohgita T, Namba N, Kono H, Shimanouchi T, Saito H. Mechanisms of enhanced aggregation and fibril formation of Parkinson's disease-related variants of α -synuclein. *Scientific reports*. 2022;12(1).1-13.
135. Tofaris GK, Goedert M, Spillantini MG. The transcellular propagation and intracellular trafficking of α -synuclein. *Cold Spring Harbor perspectives in medicine*. 2017;7(9).a024380.
136. Marques O, Outeiro TF. Alpha-synuclein: from secretion to dysfunction and death. *Cell death & disease*. 2012;3(7).e350.
137. Uemura N, Uemura MT, Luk KC, Lee VM, Trojanowski JQ. Cell-to-cell transmission of tau and α -synuclein. *Trends in molecular medicine*. 2020;26(10).936-952.
138. Perrett RM, Alexopoulou Z, Tofaris GK. The endosomal pathway in Parkinson's disease. *Molecular and Cellular Neuroscience*. 2015;66.21-28.
139. Henderson MX, Trojanowski JQ, Lee VM. α -Synuclein pathology in Parkinson's disease and related α -synucleinopathies. *Neuroscience letters*. 2019;709.134316.

140. Pieperhoff P, Südmeyer M, Dinkelbach L, Hartmann CJ, Ferrea S, Moldovan AS, Minnerop M, Diaz-Pier S, Schnitzler A, Amunts K. Regional changes of brain structure during progression of idiopathic Parkinson's disease—A longitudinal study using deformation-based morphometry. *cortex*. 2022;151:188-210.
141. Gandhi J, Antonelli AC, Afridi A, Vatsia S, Joshi G, Romanov V, Murray IV, Khan SA. Protein misfolding and aggregation in neurodegenerative diseases: a review of pathogeneses, novel detection strategies, and potential therapeutics. *Reviews in the Neurosciences*. 2019;30(4):339-358.
142. Sala G, Marinig D, Arosio A, Ferrarese C. Role of chaperone-mediated autophagy dysfunctions in the pathogenesis of Parkinson's disease. *Frontiers in molecular neuroscience*. 2016;9:57.
143. Dilsizoglu Senol A, Samarani M, Syan S, Guardia CM, Nonaka T, Liv N, Latour-Lambert P, Hasegawa M, Klumperman J, Bonifacino JS, Zurzolo C. α -Synuclein fibrils subvert lysosome structure and function for the propagation of protein misfolding between cells through tunneling nanotubes. *PLoS biology*. 2021;19(7):e3001287.
144. Decressac M, Mattsson B, Weikop P, Lundblad M, Jakobsson J, Björklund A. TFEB-mediated autophagy rescues midbrain dopamine neurons from α -synuclein toxicity. *Proceedings of the National Academy of Sciences*. 2013;110(19):E1817-1826.
145. Guiney SJ, Adlard PA, Lei P, Mawal CH, Bush AI, Finkelstein DI, Ayton S. Fibrillar α -synuclein toxicity depends on functional lysosomes. *Journal of Biological Chemistry*. 2020;295(51):17497-17513.
146. Kasen A, Houck C, Burmeister AR, Sha Q, Brundin L, Brundin P. Upregulation of α -synuclein following immune activation: Possible trigger of Parkinson's disease. *Neurobiology of Disease*. 2022:105654.
147. Rizor A, Pajarillo E, Johnson J, Aschner M, Lee E. Astrocytic oxidative/nitrosative stress contributes to Parkinson's disease pathogenesis: the dual role of reactive astrocytes. *Antioxidants*. 2019;8(8):265.
148. Minakaki G, Krainc D, Burbulla LF. The convergence of alpha-synuclein, mitochondrial, and lysosomal pathways in vulnerability of midbrain dopaminergic neurons in Parkinson's disease. *Frontiers in Cell and Developmental Biology*. 2020;8:580634.
149. Cali T, Ottolini D, Negro A, and Brini M. α -Synuclein controls mitochondrial calcium homeostasis by enhancing endoplasmic reticulum-mitochondria interactions. *Journal of Biological Chemistry*. 2012;287(22):17914-17929.

150. Di Maio R, Barrett PJ, Hoffman EK, Barrett CW, Zharikov A, Borah A, Hu X, McCoy J, Chu CT, Burton EA, Hastings TG. α -Synuclein binds to TOM20 and inhibits mitochondrial protein import in Parkinson's disease. *Science translational medicine*. 2016;8(342).342ra78.
151. Duce JA, Wong BX, Durham H, Devedjian JC, Smith DP, Devos D. Post translational changes to α -synuclein control iron and dopamine trafficking; a concept for neuron vulnerability in Parkinson's disease. *Molecular neurodegeneration*. 2017;12(1).1-12.
152. Hawk BJ, Khounlo R, Shin YK. Alpha-synuclein continues to enhance SNARE-dependent vesicle docking at exorbitant concentrations. *Frontiers in neuroscience*. 2019;13.1-10.
153. Hallaceli E, Kayatekin C, Nazeen S, Wang XH, Sheinkopf Z, Sathyakumar S, Sarkar S, Jiang X, Dong X, Di Maio R, Wang W. The Parkinson's disease protein alpha-synuclein is a modulator of processing bodies and mRNA stability. *Cell*. 2022;185(12).2035-2056.
154. Owyong TC, Hong Y. Emerging fluorescence tools for the study of proteostasis in cells. *Current Opinion in Chemical Biology*. 2022;67.102116.
155. Dharshini SAP, Jemimah S, Taguchi YH. Exploring Common Therapeutic Targets for Neurodegenerative Disorders Using Transcriptome Study. *Frontiers in genetics*. 2021;12.1-15.
156. Raj T, Li YI, Wong G, Humphrey J, Wang M, Ramdhani S, Wang YC, Ng B, Gupta I, Haroutunian V, Schadt EE. Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nature genetics*. 2018;50(11).1584-1592.
157. Pai M, Kuo Y, Wang I, Chiang P, Tsai, K. The Role of Methylated Circulating Nucleic Acids as a Potential Biomarker in Alzheimer's disease. 2019.2440-2449.
158. Kursu, M.B. Robustness of Random Forest-based gene selection methods. *BMC bioinformatics*. 2014;15(1).1-8.
159. Sarica A, Cerasa A, Quattrone A. Random Forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Frontiers in aging neuroscience*. 2017;9.329.
160. Kursu MB, Rudnicki WR. Feature Selection with the Boruta Package. *J. Stat. Softw*. 2010;36.1-13
161. Zhu Y, Stephens RM, Meltzer PS, Davis SR. SRADB: Query and use public next-generation sequencing data from within R. *BMC Bioinform*. 2013;14.19.

162. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30.2114-2120.
163. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol*. 2016;34.525-527.
164. Zhang C, Zhang B, Lin L.-L, Zhao S. Evaluation, and comparison of computational tools for RNA-seq isoform quantification. *BMC Genom*. 2017;18.583.
165. Wang T, Li B, Nelson C.E, Nabavi S. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinform*. 2019;20.40.
166. Maza E. In papyro comparison of TMM (edgeR), RLE (DESeq2), and MRN normalisation methods for a simple two-conditions-without-replicates RNA-seq experimental design. *Front. Genet*. 2016;7.164.
167. Abbas M, Yasser EM. Machine learning based refined differential gene expression analysis of pediatric sepsis. *BMC Med. Genom*. 2020;13.122.
168. Acharjee A, Larkman J, Xu Y, Cardoso VR, Gkoutos GV. A random forest-based biomarker discovery and power analysis framework for diagnostics research. *BMC Med. Genom*. 2020;13.178.
169. Stephens D, Diesing M. A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data. *PLoS ONE* 2014;9.e93950.
170. Alexa A, Rahnenführer J. Gene set enrichment analysis with topGO. *Bioconduct. Improv*. 2009;27.1-26.
171. Warde-Farley D, Donaldson S.L, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes C.T, et al. The GeneMANIA prediction server: Biological network integration for gene prioritisation and predicting gene function. *Nucleic Acids Res*. 2010;38.214-220.
172. Margiotta, A. Role of SNAREs in Neurodegenerative Diseases. *Cells*. 2021;10(5).1-15.
173. Yan C, Jiang J, Yang Y, Geng X, Dong W. The function of VAMP2 in mediating membrane fusion: An overview. *Frontiers in Molecular Neuroscience*. 2022;15.948160.
174. Huang M, Wang B, Li X, Fu C, Wang C, Kang X. α -Synuclein: a multifunctional player in exocytosis, endocytosis, and vesicle recycling. *Frontiers in Neuroscience*. 2019;13.28.
175. Haytural H, Jordà-Siquier T, Winblad B, Mulle C, Tjernberg LO, Granholm AC, Frykman S, Barthet G. Distinctive alteration of presynaptic proteins in the outer molecular layer of the dentate gyrus in Alzheimer's disease. *Brain communications*. 2021;3(2).

176. Gries M, Christmann A, Schulte S, Weyland M, Rommel S, Martin M, Baller M, Röth R, Schmitteckert S, Unger M, Liu Y. Parkinson mice show functional and molecular changes in the gut long before motoric disease onset. *Molecular neurodegeneration*. 2021;16(1).1-23.
177. Anderson KM, Augusto DG, Dandekar R, Shams H, Zhao C, Yusufali T, Montero-Martín G, Marin WM, Nemat-Gorgani N, Creary LE, Caillier S. Killer cell immunoglobulin-like receptor variants are associated with protection from symptoms associated with more severe course in Parkinson disease. *The Journal of Immunology*. 2020;205(5).1323-1330.
178. Lehmann DJ, Wiebusch H, Marshall SE, Johnston C, Warden DR, Morgan K, Schappert K, Poirier J, Xuereb J, Kalsheker N, Welsh KI. HLA class I, II & III genes in confirmed late-onset Alzheimer's disease. *Neurobiology of aging*. 2001;22(1).71-77.
179. Cifuentes RA, Murillo-Rojas J. Alzheimer's disease and HLA-A2: Linking neurodegenerative to immune processes through an in silico approach. *BioMed research international*. 2014.
180. Solana C, Tarazona R, Solana R. Immunosenescence of natural killer cells, inflammation, and Alzheimer's disease. *International Journal of Alzheimer's Disease*. 2018.1-9.
181. Liaunardy-Jopeace A, Bryant CE, Gay NJ. The COP II adaptor protein TMED7 is required to initiate and mediate the delivery of TLR4 to the plasma membrane. *Science signalling*. 2014;7(336).1-30.
182. Doyle SL, Husebye H, Connolly DJ, Espevik T, O'Neill LA, McGettrick AF. The GOLD domain-containing protein TMED7 inhibits TLR4 signalling from the endosome upon LPS stimulation. *Nature communications*. 2012; 3(1).1-11.
183. Shin JH, Park SJ, Jo DS, Park NY, Kim JB, Bae JE, Jo YK, Hwang JJ, Lee JA, Jo DG, Kim JC. Down-regulated TMED10 in Alzheimer disease induces autophagy via ATG4B activation. *Autophagy*. 2019;15(9).1495-1505.
184. Ireland SC, Huang H, Zhang J, Li J, Wang Y. Hydrogen peroxide induces Arl1 degradation and impairs Golgi-mediated trafficking. *Molecular biology of the cell*. 2020;31(17).1931-1942.
185. Lowe M. The physiological functions of the golgin vesicle tethering proteins. *Frontiers in cell and developmental biology*. 2019;7.94.

186. Li X, Yu J, Gong L, Zhang Y, Dong S, Shi J, Li C, Li Y, Zhang Y, Li H. Heme oxygenase-1 (HO-1) regulates Golgi stress and attenuates endotoxin-induced acute lung injury through hypoxia inducible factor-1 α (HIF-1 α)/HO-1 signalling pathway. *Free Radical Biology and Medicine*. 2021;165:243-253.
187. Zhang X, Wu Y, Cai F, Song W. Regulation of global gene expression in brain by TMP21. *Molecular brain*. 2019;12(1):1-10.
188. Martínez-Menárguez JÁ, Tomás M, Martínez-Martínez N, Martínez-Alonso E. Golgi fragmentation in neurodegenerative diseases: is there a common cause?. *Cells*, 2019;8(7):748.
189. Deczkowska A, Matcovitch-Natan O, Tsitsou-Kampeli A, Ben-Hamo S, Dvir-Szternfeld R, Spinrad A, Singer O, David E, Winter DR, Smith LK, Kertser A. Mef2C restrains microglial inflammatory response and is lost in brain ageing in an IFN-I-dependent manner. *Nature communications*. 2017;8(1):1-13.
190. Xue F, Tian J, Yu C, Du H, Guo L. Type I interferon response-related microglial Mef2c deregulation at the onset of Alzheimer's pathology in 5 $\times$ FAD mice. *Neurobiology of disease*. 2021;152:1-10.
191. Gu X, Fu C, Lin L, Liu S, Su X, Li A, Wu Q, Jia C, Zhang P, Chen L, Zhu X. miR-124 and miR-9 mediated downregulation of HDAC5 promotes neurite development through activating MEF2C-GPM6A pathway. *Journal of Cellular Physiology*. 2018;233(1):673-687.
192. Iannuzzi F, Sirabella R, Canu N, Maier TJ, Annunziato L, Matrone C. Fyn tyrosine kinase elicits amyloid precursor protein Tyr682 phosphorylation in neurons from Alzheimer's disease patients. *Cells*. 2020;9(8):1-21.
193. Li C, Götz J. Somatodendritic accumulation of Tau in Alzheimer's disease is promoted by Fyn-mediated local protein translation. *The EMBO journal*. 2017;36(21):3120-3138.
194. Panicker N, Saminathan H, Jin H, Neal M, Harischandra DS, Gordon R, Kanthasamy K, Lawana V, Sarkar S, Luo J, Anantharam V. Fyn kinase regulates microglial neuroinflammatory responses in cell culture and animal models of Parkinson's disease. *Journal of Neuroscience*. 2015;35(27):10058-10077.
195. Ulrich JD, Holtzman DM. TREM2 function in Alzheimer's disease and neurodegeneration. *ACS chemical neuroscience*, 2016;7(4):420-427.
196. Auburger G, Gispert S, Brehm N. Methyl-arginine profile of brain from aged PINK1-KO⁺ A53T-SNCA mice suggests altered mitochondrial biogenesis. *Parkinson's disease*. 2016.

197. Doucet-Beaupré H, Gilbert C, Profes MS, Chabrat A, Pacelli C, Giguère N, Rioux V, Charest J, Deng Q, Laguna A, Ericson J. Lmx1a and Lmx1b regulate mitochondrial functions and survival of adult midbrain dopaminergic neurons. *Proceedings of the National Academy of Sciences*. 2016;113(30).E4387-E4396.
198. Ishigaki S, Fujioka Y, Okada Y, Riku Y, Udagawa T, Honda D, Yokoi S, Endo K, Ikenaka K, Takagi S, Iguchi Y. Altered tau isoform ratio caused by loss of FUS and SFPQ function leads to FTLD-like phenotypes. *Cell reports*. 2017;18(5).1118-1131.
199. Salton M, Lerenthal Y, Wang SY, Chen DJ, Shiloh Y. Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response. *Cell cycle*. 2010;9(8).1568-1576.
200. Yoon YS, You JS, Kim TK, Ahn WJ, Kim MJ, Son KH, Ricarte D, Ortiz D, Lee SJ, Lee HJ. Senescence and impaired DNA damage responses in alpha-synucleinopathy models. *Experimental & Molecular Medicine*. 2022.1-14.
201. Sterpka A, Yang J, Strobel M, Zhou Y, Pauplis C, Chen X. Diverged morphology changes of astrocytic and neuronal primary cilia under reactive insults. *Molecular brain*, 2020;13(1).1-16.
202. Hu L, Wang B, Zhang Y. Serotonin 5-HT6 receptors affect cognition in a mouse model of Alzheimer's disease by regulating cilia function. *Alzheimer's research & therapy*, 2017;9(1).1-17.
203. Baek H, Shin HJ, Kim JJ, Shin N, Kim S, Yi MH, Zhang E, Hong J, Kang JW, Kim Y, Kim CS. Primary cilia modulate TLR4-mediated inflammatory responses in hippocampal neurons. *Journal of neuroinflammation*. 2017;14(1).1-10.
204. Nozaki S, Katoh Y, Terada M, Michisaka S, Funabashi T, Takahashi S, Kontani K, Nakayama K. Regulation of ciliary retrograde protein trafficking by the Joubert syndrome proteins ARL13B and INPP5E. *Journal of Cell Science*. 2017;130(3).563-576.
205. Swiderski RE, Nakano Y, Mullins RF, Seo S, Banfi B. A mutation in the mouse ttc26 gene leads to impaired hedgehog signalling. *PLoS genetics*, 2014;10(10).1004689.
206. Quinn EM, Cormican P, Kenny EM, Hill M, Anney R, Gill M, Corvin AP, Morris DW. Development of strategies for SNP detection in RNA-seq data: application to lymphoblastoid cell lines and evaluation using 1000 Genomes data. *PloS one*. 2013;8(3).e58815.
207. Jehl F, Degalez F, Bernard M, Lecerf F, Lagoutte L, Désert C, Coulée M, Bouchez O, Leroux S, Abasht B, Tixier-Boichard M. RNA-seq data for reliable SNP detection and genotype calling: interest for coding variant

- characterization and cis-regulation analysis by allele-specific expression in livestock species. *Frontiers in genetics*. 2021;12.1104.
- 208.Bakhtiarizadeh MR, Alamouti AA. RNA-Seq based genetic variant discovery provides new insights into controlling fat deposition in the tail of sheep. *Scientific Reports*. 2020;10(1).1-3.
- 209.Zhao Y, Wang K, Wang WL, Yin TT, Dong WQ, Xu CJ. A high-throughput SNP discovery strategy for RNA-seq data. *BMC genomics*. 2019;20(1).1-0.
- 210.Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1).15-21.
- 211.Brouard JS, Schenkel F, Marete A, Bissonnette N. The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *Journal of animal science and biotechnology*. 2019;10(1).1-6.
- 212.McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The ensembl variant effect predictor. *Genome biology*. 2016;17(1).1-4.
- 213.Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4(1).s13742-015.
- 214.Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and rigorous computation of gene and pathway scores from SNP-based summary statistics. *PLoS computational biology*. 2016;12(1).e1004714.
- 215.Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic acids research*. 2019;47(W1).199-205.
- 216.Arcos-Burgos M, Vélez JI, Martínez AF, Ribasés M, Ramos-Quiroga JA, Sánchez-Mora C, Richarte V, Roncero C, Cormand B, Fernández-Castillo N, Casas M. ADGRL3 (LPHN3) variants predict substance use disorder. *Translational psychiatry*. 2019;9(1).1-5.
- 217.McNeill RV, Palladino VS, Brunkhorst-Kanaan N, Grimm O, Reif A, Kittel-Schneider S. Expression of the adult ADHD-associated gene ADGRL3 is dysregulated by risk variants and environmental risk factors. *The World Journal of Biological Psychiatry*. 2021;22(5).335-49.
- 218.Aidaralieva NJ, Kamino K, Kimura R, Yamamoto M, Morihara T, Kazui H, Hashimoto R, Tanaka T, Kudo T, Kida T, Okuda JI. Dynamin 2 gene is a novel susceptibility gene for late-onset Alzheimer disease in non-APOE-ε4 carriers. *Journal of human genetics*. 2008;53(4).296-302.

219. Tang SS, Wang HF, Zhang W, Kong LL, Zheng ZJ, Tan MS, Tan CC, Wang ZX, Tan L, Jiang T, Yu JT. MEF2C rs190982 polymorphism with late-onset Alzheimer's disease in Han Chinese: a replication study and meta-analyses. *Oncotarget*. 2016;7(26):39136.
220. Singh M, Jadhav HR, Bhatt T. Dynamin functions and ligands: classical mechanisms behind. *Molecular pharmacology*. 2017;91(2):123-134.
221. Kamagata E, Kudo T, Kimura R, Tanimukai H, Morihara T, Sadik MG, Kamino K, Takeda M. Decrease of dynamin 2 levels in late-onset Alzheimer's disease alters A β metabolism. *Biochemical and biophysical research communications*. 2009;379(3):691-695.
222. Tanabe K, Takei K. Dynamic instability of microtubules requires dynamin 2 and is impaired in a Charcot-Marie-Tooth mutant. *Journal of cell biology*. 2009;185(6):939-948.
223. Joshi G, Wang Y. Golgi defects enhance APP amyloidogenic processing in Alzheimer's disease. *Bioessays*. 2015;37(3):240-247.
224. Xiao X, Zheng F, Chang H, Ma Y, Yao YG, Luo XJ, Li M. The gene encoding protocadherin 9 (PCDH9), a novel risk factor for major depressive disorder. *Neuropsychopharmacology*. 2018;43(5):1128-37.
225. International Parkinson's Disease Genomics Consortium (IPDGC), Wellcome Trust Case Control Consortium 2 (WTCCC2). A two-stage meta-analysis identifies several new loci for Parkinson's disease. *PLoS genetics*. 2011;7(6):1002142.
226. Iacoangeli A, Lin T, Al Khleifat A, Jones AR, Opie-Martin S, Coleman JR, Shatunov A, Sproviero W, Williams KL, Garton F, Restuadi R. Genome-wide meta-analysis finds the ACSL5-ZDHHC6 locus is associated with ALS and links weight loss to the disease genetics. *Cell reports*. 2020;33(4):108323.
227. Yoshida H, Koodie L, Jacobsen K, Hanzawa K, Miyamoto Y, Yamamoto M. B4GALNT1 induces angiogenesis, anchorage independence growth and motility, and promotes tumorigenesis in melanoma by induction of ganglioside GM2/GD2. *Scientific reports*. 2020;10(1):1-12.
228. Sipione S, Monyor J, Galleguillos D, Steinberg N, Kadam V. Gangliosides in the brain: physiology, pathophysiology and therapeutic applications. *Frontiers in neuroscience*. 2020;14:1004.
229. Chiricozzi E, Mauri L, Lunghi G, Di Biase E, Fazzari M, Maggioni M, Valsecchi M, Prioni S, Loberto N, Pomè DY, Ciampa MG. Parkinson's disease recovery by GM1 oligosaccharide treatment in the B4galnt1 $^{+/-}$ mouse model. *Scientific reports*. 2019;9(1):1-15.

- 230.Demir SA, Timur ZK, Ateş N, Martínez LA, Seyrantepe V. GM2 ganglioside accumulation causes neuroinflammation and behavioral alterations in a mouse model of early onset Tay-Sachs disease. *Journal of neuroinflammation*. 2020;17(1).1-18.
- 231.Belarbi K, Cuvelier E, Bonte MA, Desplanque M, Gressier B, Devos D, Chartier-Harlin MC. Glycosphingolipids and neuroinflammation in Parkinson's disease. *Molecular neurodegeneration*. 2020;15(1).1-16.
- 232.Zhou H, Polimanti R, Yang BZ, Wang Q, Han S, Sherva R, Nuñez YZ, Zhao H, Farrer LA, Kranzler HR, Gelernter J. Genetic risk variants associated with comorbid alcohol dependence and major depression. *JAMA psychiatry*. 2017;74(12).1234-41.
- 233.Tang Q, Schaks M, Koundinya N, Yang C, Pollard LW, Svitkina TM, Rottner K, Goode BL. WAVE1 and WAVE2 have distinct and overlapping roles in controlling actin assembly at the leading edge. *Molecular biology of the cell*, 31(20), pp.2168-2178.
- 234.Dahl JP, Wang-Dunlop J, Gonzales C, Goad ME, Mark RJ, Kwak SP. Characterization of the WAVE1 knock-out mouse: implications for CNS development. *Journal of neuroscience*. 2003;23(8).3343-3352.
- 235.Ceglia I, Reitz C, Gresack J, Ahn JH, Bustos V, Bleck M, Zhang X, Martin G, Simon SM, Nairn AC, Greengard P. APP intracellular domain–WAVE1 pathway reduces amyloid- β production. *Nature medicine*. 2015;21(9).1054-1059.
- 236.Soderling SH, Langeberg LK, Soderling JA, Davee SM, Simerly R, Raber J, Scott JD. Loss of WAVE-1 causes sensorimotor retardation and reduced learning and memory in mice. *Proceedings of the national academy of sciences*. 2003;100(4).1723-1728.
- 237.Xie T, Stathopoulou MG, De Andrés F, Siest G, Murray H, Martin M, Cobaleda J, Delgado A, Lamont J, Peñas-Lledó E, LLerena A. VEGF-related polymorphisms identified by GWAS and risk for major depression. *Translational psychiatry*. 2017;7(3).e1055-e1055.
- 238.Wu Y, Zhang Y, Han X, Li X, Xue L, Xie A. Association of VEGF gene polymorphisms with sporadic Parkinson's disease in Chinese Han population. *Neurological sciences*. 37(12), pp.1923-1929.
- 239.Lan G, Wang P, Chan RB, Liu Z, Yu Z, Liu X, Yang Y, Zhang J. Astrocytic VEGFA: An essential mediator in blood–brain-barrier disruption in Parkinson's disease. *Glia*. 2022;70(2).337-353.

240. Ali M, Bracko O. VEGF paradoxically reduces cerebral blood flow in Alzheimer's disease mice. *Neuroscience insights*. 2022;17.26331055221109254.
241. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JR, Qiao Z, Als TD, Bigdeli TB, Børte S, Bryois J, Charney AW. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature genetics*. 2021;53(6).817-829.
242. Lee HS, Lobbstaël E, Vermeire S, Sabino J, Cleynen I. Inflammatory bowel disease and Parkinson's disease: common pathophysiological links. *Gut*. 2021;70(2).408-417.
243. Parisiadou L, Bethani I, Michaki V, Krousti K, Rapti G, Efthimiopoulos S. Homer2 and Homer3 interact with amyloid precursor protein and inhibit A β production. *Neurobiology of disease*. 2008;30(3).353-364.
244. Huang GN, Huso DL, Bouyain S, Tu J, McCorkell KA, May MJ, Zhu Y, Lutz M, Collins S, Dehoff M, Kang S. NFAT binding and regulation of T cell activation by the cytoplasmic scaffolding Homer proteins. *Science*. 2008;319(5862).476-481.
245. Abdul HM, Sama MA, Furman JL, Mathis DM, Beckett TL, Weidner AM, Patel ES, Baig I, Murphy MP, LeVine H, Kraner SD. Cognitive decline in Alzheimer's disease is associated with selective changes in calcineurin/NFAT signalling. *Journal of Neuroscience*. 2009;29(41).12957-12969.
246. Weber F, Bohrmann B, Niewoehner J, Fischer JA, Rueger P, Tiefenthaler G, Moelleken J, Bujotzek A, Brady K, Singer T, Ebeling M. Brain shuttle antibody for Alzheimer's disease with attenuated peripheral effector function due to an inverted binding mode. *Cell reports*. 2018;22(1).149-162.
247. Pottier C, Zhou X, Perkerson III RB, Baker M, Jenkins GD, Serie DJ, Ghidoni R, Benussi L, Binetti G, de Munain AL, Zulaica M. Potential genetic modifiers of disease risk and age at onset in patients with frontotemporal lobar degeneration and GRN mutations: a genome-wide association study. *The lancet neurology*. 2018;17(6).548-558.
248. Wan L, Ai JQ, Yang C, Jiang J, Zhang QL, Luo ZH, Huang RJ, Tu T, Pan A, Tu E, Manavis J. Expression of the Excitatory Postsynaptic Scaffolding Protein, Shank3, in Human Brain: Effect of Age and Alzheimer's Disease. *Frontiers in aging neuroscience*. 2021.534.
249. Delling JP, Boeckers TM. Comparison of SHANK3 deficiency in animal models: Phenotypes, treatment strategies, and translational implications. *Journal of neurodevelopmental disorders*. 2021;13(1).1-37.

- 250.Lin R, Learman LN, Bangash MA, Melnikova T, Leyder E, Reddy SC, Naidoo N, Park JM, Savonenko A, Worley PF. Homer1a regulates Shank3 expression and underlies behavioral vulnerability to stress in a model of Phelan-McDermid syndrome. *Cell reports*. 2021;37(7).110014.
- 251.Mossa A, Pagano J, Ponzoni L, Tozzi A, Vezzoli E, Sciacaluga M, Costa C, Beretta S, Francolini M, Sala M, Calabresi P. Developmental impaired Akt signalling in the Shank1 and Shank3 double knock-out mice. *Molecular psychiatry*. 2021;26(6).1928-1944.
- 252.Mills KH. IL-17 and IL-17-producing cells in protection versus pathology. *Nature Reviews Immunology*. 2022.1-7.
- 253.Gao J, Wang Y, Ma S, Xu S, Chen M. Secukinumab alleviates cognitive impairment by attenuating oxidative stress and neuronal apoptosis via the IL-17RA/AKT/ERK1/2 pathway in a rat model of sepsis. *Experimental Neurology*. 2022.114263.
- 254.Kannarkat GT, Cook DA, Lee JK, Chang J, Chung J, Sandy E, Paul KC, Ritz B, Bronstein J, Factor SA, Boss JM. Common genetic variant association with altered HLA expression, synergy with pyrethroid exposure, and risk for Parkinson's disease: an observational and case-control study. *npj Parkinson's disease*. 2015;1(1).1-9.
- 255.Yu E, Ambati A, Andersen MS, Krohn L, Estiar MA, Saini P, Senkevich K, Sosero YL, Sreelatha AA, Ruskey JA, Asayesh F. Fine mapping of the HLA locus in Parkinson's disease in Europeans. *npj Parkinson's disease*. 2021;7(1).1-7.
- 256.Hiltunen AE, Kangas SM, Ohlmeier S, Pietilä I, Hiltunen J, Tanila H, McKerlie C, Govindan S, Tuominen H, Kaarteenaho R, Hallman M. Variant in NHLRC2 leads to increased hnRNP C2 in developing neurons and the hippocampus of a mouse model of FINCA disease. *Molecular medicine*. 2020;26(1).1-6.
- 257.Li Y, Zhang Y, Jiang G, Wang Y, He C, Zhao X, Liu L, Li L. Case report: novel mutations of NDUFS6 and NHLRC2 genes potentially cause the quick postnatal death of a Chinese Hani minority neonate with mitochondrial complex I deficiency and FINCA syndrome. *Medicine*. 2022;101(27).
- 258.Nguyen JA, Yates RM. Better together: current insights into phagosome-lysosome fusion. *Frontiers in immunology*. 2021;12.636078.

259. Tang Q, Gao P, Arzberger T, Höllerhage M, Herms J, Höglinger G, Koeglsperger T. Alpha-Synuclein defects autophagy by impairing SNAP29-mediated autophagosome-lysosome fusion. *Cell death & disease*. 2021;12(10).1-16.
260. Zafar F, Valappil RA, Kim S, Johansen KK, Chang AL, Tetrad JW, Eis PS, Hatchwell E, Langston JW, Dickson DW, Schüle B. Genetic fine-mapping of the Iowan SNCA gene triplication in a patient with Parkinson's disease. *npj Parkinson's disease*. 2018;4(1).1-7.
261. Xie T, Deng L, Mei P, Zhou Y, Wang B, Zhang J, Lin J, Wei Y, Zhang X, Xu R. A genome-wide association study combining pathway analysis for typical sporadic amyotrophic lateral sclerosis in Chinese Han populations. *Neurobiology of aging*. 2014;35(7).1778-e9.
262. Costa AS, Agostini S, Guerini FR, Mancuso R, Clerici M, Pandey JP. Relation between FCGR1B rs1050501 and HSV-1 specific IgG antibodies in Alzheimer's disease. *Journal of Translational Medicine*. 2020;18(1).1-6.
263. Manocha GD, Ghatak A, Puig KL, Kraner SD, Norris CM, Combs CK. NFATc2 modulates microglial activation in the A β PP/PS1 mouse model of Alzheimer's disease. *Journal of Alzheimer's disease*. 2017;58(3).775-87.
264. Kam TI, Park H, Gwon Y, Song S, Kim SH, Moon SW, Jo DG, Jung YK. Fc γ RIIb-SHIP2 axis links A β to tau pathology by disrupting phosphoinositide metabolism in Alzheimer's disease model. *Elife*. 2016;5.e18691.
265. Kam TI, Song S, Gwon Y, Park H, Yan JJ, Im I, Choi JW, Choi TY, Kim J, Song DK, Takai T. Fc γ RIIb mediates amyloid- β neurotoxicity and memory impairment in Alzheimer's disease. *The Journal of clinical investigation*. 2013;123(7).2791-802.
266. Patel D, Zhang X, Farrell JJ, Chung J, Stein TD, Lunetta KL, Farrer LA. Cell-type-specific expression quantitative trait loci associated with Alzheimer's disease in blood and brain tissue. *Translational Psychiatry*. 2021;11(1).1-7.
267. Steele NZ, Carr JS, Bonham LW, Geier EG, Damotte V, Miller ZA, Desikan RS, Boehme KL, Mukherjee S, Crane PK, Kauwe JS. Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer's disease: a case-control study. *PLoS medicine*. 2017;14(3).e1002272.

268. Lu RC, Yang W, Tan L, Sun FR, Tan MS, Zhang W, Wang HF, Tan L. Association of HLA-DRB1 polymorphism with Alzheimer's disease: a replication and meta-analysis. *Oncotarget*. 2017;8(54):93219.
269. Huang J, Stein TD, Wang Y, Ang TF, Tao Q, Lunetta KL, Massaro J, Akhter-Khan SC, Mez J, Au R, Farrer LA. Blood levels of MCP-1 modulate the genetic risks of Alzheimer's disease mediated by HLA-DRB1 and APOE for Alzheimer's disease. *Alzheimer's & Dementia*. 2022.
270. Jing H, Zhu JX, Wang HF, Zhang W, Zheng ZJ, Kong LL, Tan CC, Wang ZX, Tan L, Tan L. INPP5D rs35349669 polymorphism with late-onset Alzheimer's disease: a replication study and meta-analysis. *Oncotarget*. 2016;7(43):69225.
271. Tsai AP, Lin PB, Dong C, Moutinho M, Casali BT, Liu Y, Lamb BT, Landreth GE, Oblak AL, Nho K. INPP5D expression is associated with risk for Alzheimer's disease and induced by plaque-associated microglia. *Neurobiology of disease*. 2021;153:105303.
272. Waugh MG. PIPs in neurological diseases. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*. 2015;1851(8):1066-1082.
273. Volta M, Lavdas AA, Obergasteiger J, Überbacher C, Picard A, Pramstaller PP, Hicks AA, Corti C. Elevated levels of alpha-synuclein blunt cellular signal transduction downstream of Gq protein-coupled receptors. *Cellular signalling*. 2017;30:82-91.
274. Bohush A, Niewiadomska G, Filipek A. Role of mitogen activated protein kinase signalling in Parkinson's disease. *International journal of molecular sciences*. 2018;19(10):2973.
275. Kim Y, Sung JY, Ceglia I, Lee KW, Ahn JH, Halford JM, Kim AM, Kwak SP, Park JB, Ho Ryu S, Schenck A. Phosphorylation of WAVE1 regulates actin polymerization and dendritic spine morphology. *Nature*. 2006;442(7104):814-7.
276. Tomizawa K, Sunada S, Lu YF, Oda Y, Kinuta M, Ohshima T, Saito T, Wei FY, Matsushita M, Li ST, Tsutsui K. Cophosphorylation of amphiphysin I and dynamin I by Cdk5 regulates clathrin-mediated endocytosis of synaptic vesicles. *The Journal of cell biology*. 2003;163(4):813-824.
277. Gong X, Tang X, Wiedmann M, Wang X, Peng J, Zheng D, Blair LA, Marshall J, Mao Z. Cdk5-mediated inhibition of the protective effects of transcription factor MEF2 in neurotoxicity-induced apoptosis. *Neuron*. 2003;38(1):33-46.

- 278.Dunn AR, O'Connell KM, Kaczorowski CC. Gene-by-environment interactions in Alzheimer's disease and Parkinson's disease. *Neuroscience & Biobehavioral Reviews*. 2019;103:73-80.
- 279.Guo P, Gong W, Li Y, Liu L, Yan R, Wang Y, Zhang Y, Yuan Z. Pinpointing novel risk loci for Lewy body dementia and the shared genetic etiology with Alzheimer's disease and Parkinson's disease: a large-scale multi-trait association analysis. *BMC medicine*. 2022;20(1):1-6.
- 280.Liu G, Bao X, Jiang Y, Liao M, Jiang Q, Feng R, Zhang L, Ma G, Chen Z, Wang G, Wang R. Identifying the association between Alzheimer's disease and Parkinson's disease using genome-wide association studies and protein-protein interaction network. *Molecular neurobiology*. 2015;52(3):1629-36.
- 281.Park E, Pan Z, Zhang Z, Lin L, Xing Y. The expanding landscape of alternative splicing variation in human populations. *The American Journal of Human Genetics*. 2018;102(1):11-26.
- 282.Brown KE, Kelly JK. Genome-wide association mapping of transcriptome variation in *Mimulus guttatus* indicates differing patterns of selection on cis-versus trans-acting mutations. *Genetics*. 2022;220(1):iyab189.
- 283.Chao Y, Jiang Y, Zhong M, Wei K, Hu C, Qin Y, Zuo Y, Yang L, Shen Z, Zou C. Regulatory roles and mechanisms of alternative RNA splicing in adipogenesis and human metabolic health. *Cell & Bioscience*. 2021;11(1):1-6.
- 284.Titus MB, Chang AW, Olesnick EC. Exploring the diverse functional and regulatory consequences of alternative splicing in development and disease. *Frontiers in Genetics*. 2021:2380.
- 285.Harvey SE, Cheng C. Methods for characterization of alternative RNA splicing. In *Long non-coding RNAs* 2016. 229-241. Humana Press, New York, NY.
- 286.Mucaki EJ, Shirley BC, Rogan PK. Expression changes confirm genomic variants predicted to result in allele-specific, alternative mRNA splicing. *Frontiers in genetics*. 2020;11:109.
- 287.Zhang Y, Yang HT, Kadash-Edmondson K, Pan Y, Pan Z, Davidson BL, Xing Y. Regional variation of splicing QTLs in human brain. *The American Journal of Human Genetics*. 2020;107(2):196-210.
- 288.Macauda A, Piredda C, Clay-Gilmour AI, Sainz J, Buda G, Markiewicz M, Barington T, Ziv E, Hildebrandt MA, Belachew AA, Varkonyi J. Expression quantitative trait loci of genes predicting outcome are associated with survival of multiple myeloma patients. *International journal of cancer*. 2021;149(2):327-36.
- 289.Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG. Genetic analysis of genome-wide variation in human gene expression. *Nature*. 2004;430(7001):743-7.

290. Andrews SJ, Fulton-Howard B, Goate A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *The Lancet Neurology*. 2020;19(4):326-35.
291. Trincado JL, Entizne JC, Hysenaj G, Singh B, Skalic M, Elliott DJ, Eyras E. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome biology*. 2018;19(1):1-11.
292. Shabalin AA. Matrix eQTL: ultra-fast eQTL analysis via large matrix operations. *Bioinformatics*. 2012;28(10):1353-1358.
293. Garrido-Martín D, Borsari B, Calvo M, Reverter F, Guigó R. Identification, and analysis of splicing quantitative trait loci across multiple tissues in the human genome. *Nature communications*. 2021;12(1):1-6.
294. Gerst JE. SNARE regulators: matchmakers and matchbreakers. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 2003;1641(2-3):99-110.
295. Tang SJ, Fesharaki-Zadeh A, Takahashi H, Nies SH, Smith LM, Luo A, Chyung A, Chiasseu M, Strittmatter SM. Fyn kinase inhibition reduces protein aggregation, increases synapse density, and improves memory in transgenic and traumatic Tauopathy. *Acta neuropathologica communications*. 2020;8(1):1-21.
296. Zhao Y, Jaber VR, LeBeauf A, Sharfman NM, Lukiw WJ. microRNA-34a (miRNA-34a) mediated down-regulation of the post-synaptic cytoskeletal element SHANK3 in sporadic Alzheimer's disease (AD). *Frontiers in Neurology*. 2019;10:28.
297. Milovanovic J, Arsenijevic A, Stojanovic B, Kanjevac T, Arsenijevic D, Radosavljevic G, Milovanovic M, Arsenijevic N. Interleukin-17 in chronic inflammatory neurological diseases. *Frontiers in immunology*. 2020;11:947.
298. Ali M, Bracko O. VEGF Paradoxically Reduces Cerebral Blood Flow in Alzheimer's Disease Mice. *Neuroscience Insights*. 2022;17:26331055221109254.
299. Mahoney ER, Dumitrescu L, Moore AM, Cambronero FE, De Jager PL, Koran ME, Petyuk VA, Robinson RA, Goyal S, Schneider JA, Bennett DA. Brain expression of the vascular endothelial growth factor gene family in cognitive aging and alzheimer's disease. *Molecular psychiatry*. 2021;26(3):888-896.
300. Kyratzi E, Efthimiopoulos S. Calcium regulates the interaction of amyloid precursor protein with Homer3 protein. *Neurobiology of aging*. 2014;35(9):2053-63.

- 301.Morgan AW, Robinson JI, Barrett JH, Martin J, Walker A, Babbage SJ, Ollier WE, Gonzalez-Gay MA, Isaacs JD. Association of FCGR2A and FCGR2A-FCGR3A haplotypes with susceptibility to giant cell arteritis. *Arthritis research & therapy*. 2006;8(4).1-6.
- 302.Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hägg S, Athanasiu L, Voyle N. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature genetics*. 2019;51(3).404-13.
- 303.Schechter M, Sharon R. An Emerging Role for Phosphoinositides in the Pathophysiology of Parkinson's Disease. *Journal of Parkinson's Disease*. 2021 (Preprint).1-25.

Appendices

Appendix 1 Table contains information of samples downloaded from National Centre of Biotechnological Information's Sequence Read Archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>).

The SRA publicly accessioned objects available in the meta-data: Project (study): describes the goal of the research effort. SID (SRR#): a manifest of data file(s) that are derived from sequencing a library. Strategy: describes the sequencing strategy. Layout: implies the paired-end and single-end RNA-sequencing per each sample.

Type of tissue and the related human disease can be found in the Tissue and Condition column, respectively. Related physiological traits can be found in Age, Sex, Braak, RIN, and Dementia column.

Project	SID	Strategy	Layout	Condition	Tissue	Sex	Age	Braak	RIN	Dementia
PRJNA232671	SRR1063418	RNA-seq	Single	AD	DLPFC/BA9		100	4		
PRJNA232671	SRR1063419	RNA-seq	Single				88	5		
PRJNA232671	SRR1063420	RNA-seq	Single				85	5		
PRJNA210613	SRR1103937	RNA-seq	Single			1	80	2		
PRJNA210613	SRR1103938	RNA-seq	Single			1	78	2		
PRJNA210613	SRR1103939	RNA-seq	Single			2	52	2		
PRJNA210613	SRR1103940	RNA-seq	Single			2	74	2		
PRJNA210613	SRR1103941	RNA-seq	Single			1	83	2		
PRJNA210613	SRR1103942	RNA-seq	Single			2	78	2		
PRJNA210613	SRR1103943	RNA-seq	Single			1	79	5		
PRJNA210613	SRR1103944	RNA-seq	Single			2	80	6		
PRJNA210613	SRR1103945	RNA-seq	Single			1	67	6		
PRJNA210613	SRR1103946	RNA-seq	Single			2	67	7		
PRJNA210613	SRR1103947	RNA-seq	Single			2	68	8		
PRJNA210613	SRR1103948	RNA-seq	Single			2	72	9		
PRJNA232669	SRR2422926	RNA-seq	Paired							
PRJNA232669	SRR2422927	RNA-seq	Paired							
PRJNA232669	SRR2422928	RNA-seq	Paired							
PRJNA232669	SRR2422929	RNA-seq	Paired							
PRJNA232669	SRR2422930	RNA-seq	Paired							
PRJNA232669	SRR2422931	RNA-seq	Paired							
PRJNA232669	SRR2422932	RNA-seq	Paired							
PRJNA232669	SRR2422933	RNA-seq	Paired							
PRJNA232669	SRR2422934	RNA-seq	Paired							
PRJNA232669	SRR2422948	RNA-seq	Paired							
PRJNA232669	SRR2422949	RNA-seq	Paired							
PRJNA232669	SRR2422950	RNA-seq	Paired							
PRJNA232669	SRR2422951	RNA-seq	Paired							

PRJNA232669	SRR2422952	RNA-seq	Paired						
PRJNA232669	SRR2422953	RNA-seq	Paired						
PRJNA434441	SRR6742935	RNA-seq	Single			2	72	2	
PRJNA434441	SRR6742936	RNA-seq	Single			2	72	2	
PRJNA434441	SRR6742939	RNA-seq	Single			1	65	1	
PRJNA434441	SRR6742940	RNA-seq	Single			1	65	1	
PRJNA434441	SRR6742945	RNA-seq	Single			1	85	2	
PRJNA434441	SRR6742946	RNA-seq	Single			1	85	2	
PRJNA434441	SRR6742947	RNA-seq	Single			1	91	2	
PRJNA434441	SRR6742948	RNA-seq	Single			1	91	2	
PRJNA434441	SRR6742949	RNA-seq	Single			1	96	2	
PRJNA434441	SRR6742950	RNA-seq	Single			1	96	2	
PRJNA434441	SRR6742953	RNA-seq	Single			2	96	2	
PRJNA434441	SRR6742954	RNA-seq	Single			2	96	2	
PRJNA434441	SRR6742957	RNA-seq	Single			2	87	2	
PRJNA434441	SRR6742958	RNA-seq	Single			2	87	2	
PRJNA434441	SRR6742963	RNA-seq	Single			2	89	2	
PRJNA434441	SRR6742964	RNA-seq	Single			2	89	2	
PRJNA434441	SRR6742971	RNA-seq	Single			1	90	2	
PRJNA434441	SRR6742972	RNA-seq	Single			1	90	2	
PRJNA434441	SRR6742977	RNA-seq	Single			1	80	2	
PRJNA434441	SRR6742978	RNA-seq	Single			1	80	2	
PRJNA434441	SRR6742933	RNA-seq	Single			1	88	5	
PRJNA434441	SRR6742934	RNA-seq	Single			1	88	5	
PRJNA434441	SRR6742937	RNA-seq	Single			1	85	6	
PRJNA434441	SRR6742938	RNA-seq	Single			1	85	6	
PRJNA434441	SRR6742967	RNA-seq	Single			2	73	6	
PRJNA434441	SRR6742968	RNA-seq	Single			2	73	6	
PRJNA434441	SRR6742973	RNA-seq	Single			2	96	6	
PRJNA434441	SRR6742974	RNA-seq	Single			2	96	6	
PRJNA434441	SRR6742975	RNA-seq	Single			2	85	6	
PRJNA434441	SRR6742976	RNA-seq	Single			2	85	6	
PRJNA434441	SRR6742981	RNA-seq	Single			2	89	6	
PRJNA434441	SRR6742982	RNA-seq	Single			2	89	6	
PRJNA434441	SRR6742941	RNA-seq	Single			1	98	4	
PRJNA434441	SRR6742942	RNA-seq	Single			1	98	4	
PRJNA434441	SRR6742943	RNA-seq	Single			1	92	3	
PRJNA434441	SRR6742944	RNA-seq	Single			1	92	3	
PRJNA434441	SRR6742951	RNA-seq	Single			1	93	4	
PRJNA434441	SRR6742952	RNA-seq	Single			1	93	4	
PRJNA434441	SRR6742955	RNA-seq	Single			1	86	4	

PRJNA434441	SRR6742956	RNA-seq	Single			1	86	4		
PRJNA434441	SRR6742959	RNA-seq	Single			2	88	3		
PRJNA434441	SRR6742960	RNA-seq	Single			2	88	3		
PRJNA434441	SRR6742961	RNA-seq	Single			1	78	3		
PRJNA434441	SRR6742962	RNA-seq	Single			1	78	3		
PRJNA434441	SRR6742965	RNA-seq	Single			2	90	4		
PRJNA434441	SRR6742966	RNA-seq	Single			2	90	4		
PRJNA434441	SRR6742969	RNA-seq	Single			2	75	3		
PRJNA434441	SRR6742970	RNA-seq	Single			2	75	3		
PRJNA434441	SRR6742979	RNA-seq	Single			2	86	3		
PRJNA434441	SRR6742980	RNA-seq	Single			2	86	3		
PRJNA530780	SRR8847556	RNA-seq	Paired			2	79	6	7	
PRJNA530780	SRR8847557	RNA-seq	Paired			2	79	6	7	
PRJNA530780	SRR8847558	RNA-seq	Paired			1	81	6	6.5	
PRJNA530780	SRR8847560	RNA-seq	Paired			1	57	6	6.8	
PRJNA530780	SRR8847561	RNA-seq	Paired			1	57	6	6.8	
PRJNA530780	SRR8847562	RNA-seq	Paired			2	81	6	6.1	
PRJNA530780	SRR8847563	RNA-seq	Paired			2	81	6	6.1	
PRJNA530780	SRR8847566	RNA-seq	Paired			2	89	6	6.5	
PRJNA530780	SRR8847567	RNA-seq	Paired			2	89	6	6.5	
PRJNA530780	SRR8847568	RNA-seq	Paired			1	62	6	6.6	
PRJNA530780	SRR8847570	RNA-seq	Paired			2	73	6	5.7	
PRJNA530780	SRR8847571	RNA-seq	Paired			2	73	6	5.7	
PRJNA530780	SRR8847572	RNA-seq	Paired			2	73	6	5.7	
PRJNA641912	SRR12092014	ncRNA-seq	Single			2	84	5		
PRJNA641912	SRR12092015	ncRNA-seq	Single			1	78	6		
PRJNA641912	SRR12092016	ncRNA-seq	Single			2	98	5		
PRJNA641912	SRR12092017	ncRNA-seq	Single			1	86	5		
PRJNA641912	SRR12092018	ncRNA-seq	Single			1	95	4		
PRJNA641912	SRR12092019	ncRNA-seq	Single			1	76	5		
PRJNA662923	SRR12623869	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623870	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623871	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623872	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623873	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623874	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623875	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623876	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623877	RNA-seq	Paired			2	70	5		
PRJNA662923	SRR12623878	RNA-seq	Paired			2	70	5		
PRJNA662923	SRR12623879	RNA-seq	Paired			2	70	5		

PRJNA662923	SRR12623880	RNA-seq	Paired			2	70	5		
PRJNA662923	SRR12623881	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623882	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623883	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623884	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623885	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623886	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623887	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623888	RNA-seq	Paired			1	73	6		
PRJNA662923	SRR12623889	RNA-seq	Paired			2	71	6		
PRJNA662923	SRR12623890	RNA-seq	Paired			2	71	6		
PRJNA662923	SRR12623891	RNA-seq	Paired			2	71	6		
PRJNA662923	SRR12623892	RNA-seq	Paired			2	71	6		
PRJNA662923	SRR12623893	RNA-seq	Paired			1	72	5		
PRJNA662923	SRR12623894	RNA-seq	Paired			1	72	5		
PRJNA662923	SRR12623895	RNA-seq	Paired			1	72	5		
PRJNA662923	SRR12623896	RNA-seq	Paired			1	72	5		
PRJNA662923	SRR12623897	RNA-seq	Paired			1	69	5		
PRJNA662923	SRR12623898	RNA-seq	Paired			1	69	5		
PRJNA662923	SRR12623899	RNA-seq	Paired			1	69	5		
PRJNA662923	SRR12623900	RNA-seq	Paired			1	69	5		
PRJNA662923	SRR12623901	RNA-seq	Paired			2	70	4		
PRJNA662923	SRR12623902	RNA-seq	Paired			2	70	4		
PRJNA662923	SRR12623903	RNA-seq	Paired			2	70	4		
PRJNA662923	SRR12623904	RNA-seq	Paired			2	70	4		
PRJNA662923	SRR12623905	RNA-seq	Paired			1	73	5		
PRJNA662923	SRR12623906	RNA-seq	Paired			1	73	5		
PRJNA662923	SRR12623907	RNA-seq	Paired			1	73	5		
PRJNA662923	SRR12623908	RNA-seq	Paired			1	73	5		
PRJNA662923	SRR12623909	RNA-seq	Paired			1	72	6		
PRJNA662923	SRR12623910	RNA-seq	Paired			1	72	6		
PRJNA662923	SRR12623911	RNA-seq	Paired			1	72	6		
PRJNA662923	SRR12623912	RNA-seq	Paired			1	72	6		
PRJNA662923	SRR12623913	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623914	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623915	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623916	RNA-seq	Paired			1	74	4		
PRJNA662923	SRR12623917	RNA-seq	Paired			1	69	4		
PRJNA662923	SRR12623918	RNA-seq	Paired			1	69	4		
PRJNA662923	SRR12623919	RNA-seq	Paired			1	69	4		
PRJNA662923	SRR12623920	RNA-seq	Paired			1	69	4		

PRJNA662923	SRR12623921	RNA-seq	Paired			2	70	5		
PRJNA662923	SRR12623922	RNA-seq	Paired			2	70	5		
PRJNA662923	SRR12623923	RNA-seq	Paired			2	70	5		
PRJNA662923	SRR12623924	RNA-seq	Paired			2	70	5		
PRJEB5212	ERR409878	RNA-seq	Single	Ctrl	DLPFC/BA9	1	43			
PRJEB5212	ERR409879	RNA-seq	Single			1	80			
PRJEB5212	ERR409882	RNA-seq	Single			1	53			
PRJEB5212	ERR409883	RNA-seq	Single			1	46			
PRJEB5212	ERR409887	RNA-seq	Single			1	76			
PRJEB5212	ERR409888	RNA-seq	Single			1	57			
PRJEB5212	ERR409891	RNA-seq	Single			1	79			
PRJEB5212	ERR409893	RNA-seq	Single			1	40			
PRJEB5212	ERR409896	RNA-seq	Single			1	57			
PRJEB5212	ERR409898	RNA-seq	Single			1	93			
PRJEB5212	ERR409900	RNA-seq	Single			1	44			
PRJNA143627	SRR306840	RNA-seq	Single			1			8.6	
PRJNA143627	SRR306841	RNA-seq	Single			1			8.5	
PRJNA143627	SRR306842	RNA-seq	Single			1			7.8	
PRJNA185476	SRR646457	RNA-seq	Single			2				
PRJNA185476	SRR646458	RNA-seq	Single			2				
PRJNA185476	SRR646459	RNA-seq	Single			2				
PRJNA185476	SRR646460	RNA-seq	Single			2				
PRJNA232671	SRR1063412	RNA-seq	Single				85	0		
PRJNA232671	SRR1063413	RNA-seq	Single				80	1		
PRJNA232671	SRR1063414	RNA-seq	Single				90	0		
PRJNA232669	SRR2422918	RNA-seq	Paired							
PRJNA232669	SRR2422919	RNA-seq	Paired							
PRJNA232669	SRR2422920	RNA-seq	Paired							
PRJNA232669	SRR2422922	RNA-seq	Paired							
PRJNA232669	SRR2422923	RNA-seq	Paired							
PRJNA232669	SRR2422924	RNA-seq	Paired							
PRJNA232669	SRR2422925	RNA-seq	Paired							
PRJNA232669	SRR2422943	RNA-seq	Paired							
PRJNA232669	SRR2422944	RNA-seq	Paired							
PRJNA232669	SRR2422945	RNA-seq	Paired							
PRJNA232669	SRR2422946	RNA-seq	Paired							
PRJNA232669	SRR2422947	RNA-seq	Paired							
PRJNA641912	SRR12092020	ncRNA-seq	Single			1	85	2		
PRJNA641912	SRR12092021	ncRNA-seq	Single			2	74	1		
PRJNA641912	SRR12092022	ncRNA-seq	Single			2	79	2		
PRJNA641912	SRR12092023	ncRNA-seq	Single			1	86	2		

PRJNA662923	SRR12623925	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623926	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623927	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623928	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623929	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623930	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623931	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623932	RNA-seq	Paired			1	85	1		
PRJNA662923	SRR12623933	RNA-seq	Paired			2	80	1		
PRJNA662923	SRR12623934	RNA-seq	Paired			2	80	1		
PRJNA662923	SRR12623935	RNA-seq	Paired			2	80	1		
PRJNA662923	SRR12623936	RNA-seq	Paired			2	80	1		
PRJNA662923	SRR12623937	RNA-seq	Paired			1	81	1		
PRJNA662923	SRR12623938	RNA-seq	Paired			1	81	1		
PRJNA662923	SRR12623939	RNA-seq	Paired			1	81	1		
PRJNA662923	SRR12623940	RNA-seq	Paired			1	81	1		
PRJNA662923	SRR12623941	RNA-seq	Paired			2	79	1		
PRJNA662923	SRR12623942	RNA-seq	Paired			2	79	1		
PRJNA662923	SRR12623943	RNA-seq	Paired			2	79	1		
PRJNA662923	SRR12623944	RNA-seq	Paired			2	79	1		
PRJNA662923	SRR12623945	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623946	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623947	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623948	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623949	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623950	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623951	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623952	RNA-seq	Paired			1	76	1		
PRJNA662923	SRR12623953	RNA-seq	Paired			2	82	1		
PRJNA662923	SRR12623954	RNA-seq	Paired			2	82	1		
PRJNA662923	SRR12623955	RNA-seq	Paired			2	82	1		
PRJNA662923	SRR12623956	RNA-seq	Paired			2	82	1		
PRJNA662923	SRR12623957	RNA-seq	Paired			1	84	1		
PRJNA662923	SRR12623958	RNA-seq	Paired			1	84	1		
PRJNA662923	SRR12623959	RNA-seq	Paired			1	84	1		
PRJNA662923	SRR12623960	RNA-seq	Paired			1	84	1		
PRJNA174763	SRR563541	RNA-seq	Single	PD	DLPFC/BA9	1	78			
PRJNA174763	SRR563542	RNA-seq	Single			1	73			
PRJNA174763	SRR563543	RNA-seq	Single			1	83			
PRJNA174763	SRR563544	RNA-seq	Single			1	80			
PRJNA174763	SRR563545	RNA-seq	Single			1	74			

PRJNA174763	SRR563546	RNA-seq	Single			1	76			
PRJNA174763	SRR563547	RNA-seq	Single			2	82			
PRJNA174763	SRR563548	RNA-seq	Single			1	81			
PRJNA174763	SRR563549	RNA-seq	Single			2	77			
PRJNA174763	SRR563550	RNA-seq	Single			1	74			
PRJNA174763	SRR563551	RNA-seq	Single			1	76			
PRJNA174763	SRR563552	RNA-seq	Single			2	78			
PRJNA174763	SRR563553	RNA-seq	Single			1	79			
PRJNA174763	SRR563554	RNA-seq	Single			2	83			
PRJNA174763	SRR563555	RNA-seq	Single			1	75			
PRJNA174763	SRR563556	RNA-seq	Single			1	77			
PRJNA174763	SRR563557	RNA-seq	Single			2	82			
PRJNA283498	SRR2015757	RNA-seq	Paired			1	74	1	7.2	1
PRJNA283498	SRR2015758	RNA-seq	Paired			1	83	3	7.1	1
PRJNA283498	SRR2015759	RNA-seq	Paired			1	80		8.5	0
PRJNA283498	SRR2015760	RNA-seq	Paired			1	83	3	7.3	1
PRJNA283498	SRR2015761	RNA-seq	Paired			1	80	2	8.4	0
PRJNA283498	SRR2015762	RNA-seq	Paired			1	84	2	6.7	1
PRJNA283498	SRR2015763	RNA-seq	Paired			1	88	2	8.1	1
PRJNA283498	SRR2015764	RNA-seq	Paired			1	81	1	5.8	0
PRJNA283498	SRR2015765	RNA-seq	Paired			1	77	3	6.1	0
PRJNA283498	SRR2015766	RNA-seq	Paired			1	64	2	8.1	0
PRJNA283498	SRR2015767	RNA-seq	Paired			1	85	3	6.1	1
PRJNA283498	SRR2015768	RNA-seq	Paired			1	94	3	6.4	
PRJNA283498	SRR2015769	RNA-seq	Paired			1	80	1	6.5	1
PRJNA283498	SRR2015770	RNA-seq	Paired			1	85	1	6.9	0
PRJNA283498	SRR2015771	RNA-seq	Paired			1	75	1	8	
PRJNA283498	SRR2015772	RNA-seq	Paired			1	74	2	7.1	
PRJNA283498	SRR2015773	RNA-seq	Paired			1	89	2	6.1	0
PRJNA283498	SRR2015774	RNA-seq	Paired			1	66	1	7.7	0
PRJNA283498	SRR2015775	RNA-seq	Paired			1	65	1	6.8	
PRJNA283498	SRR2015776	RNA-seq	Paired			1	85		6.1	
PRJNA283498	SRR2015777	RNA-seq	Paired			1	64	2	7.8	1
PRJNA283498	SRR2015778	RNA-seq	Paired			1	64	1	6.1	0
PRJNA283498	SRR2015779	RNA-seq	Paired			1	75	2	6.9	1
PRJNA283498	SRR2015780	RNA-seq	Paired			1	68	1	7.3	1
PRJNA283498	SRR2015781	RNA-seq	Paired			1	95	1	7	1
PRJNA283498	SRR2015782	RNA-seq	Paired			1	74	2	7.5	0
PRJNA283498	SRR2015783	RNA-seq	Paired			1	68	2	7.9	
PRJNA283498	SRR2015784	RNA-seq	Paired			1	79	2	6.9	0
PRJNA283498	SRR2015785	RNA-seq	Paired			1	70	3	6.7	0

PRJNA290524	SRR2121909	RNA-seq	Paired								
PRJNA295431	SRR2353417	miRNA-seq	Single				1	74		7.7	1
PRJNA295431	SRR2353418	miRNA-seq	Single				1	83		6.85	1
PRJNA295431	SRR2353419	miRNA-seq	Single				1	80		8.5	0
PRJNA295431	SRR2353420	miRNA-seq	Single				1	83		8.05	1
PRJNA295431	SRR2353421	miRNA-seq	Single				1	80		7.7	0
PRJNA295431	SRR2353422	miRNA-seq	Single				1	84		7.2	1
PRJNA295431	SRR2353423	miRNA-seq	Single				1	88		7.5	1
PRJNA295431	SRR2353424	miRNA-seq	Single				1	81		6.4	0
PRJNA295431	SRR2353425	miRNA-seq	Single				1	77		7.2	0
PRJNA295431	SRR2353426	miRNA-seq	Single				1	64		8	0
PRJNA295431	SRR2353427	miRNA-seq	Single				1	85		8.2	1
PRJNA295431	SRR2353428	miRNA-seq	Single				1	94		7.1	0
PRJNA295431	SRR2353429	miRNA-seq	Single				1	80		7.9	1
PRJNA295431	SRR2353430	miRNA-seq	Single				1	85		7.2	0
PRJNA295431	SRR2353431	miRNA-seq	Single				1	75		8	0
PRJNA295431	SRR2353432	miRNA-seq	Single				1	74		6.2	0
PRJNA295431	SRR2353433	miRNA-seq	Single				1	89		6.8	0
PRJNA295431	SRR2353434	miRNA-seq	Single				1	66		8.2	0
PRJNA295431	SRR2353435	miRNA-seq	Single				1	65		8	0
PRJNA295431	SRR2353436	miRNA-seq	Single				1	85		5.5	0
PRJNA295431	SRR2353437	miRNA-seq	Single				1	64		7.8	1
PRJNA295431	SRR2353438	miRNA-seq	Single				1	64		6.1	0
PRJNA295431	SRR2353439	miRNA-seq	Single				1	75		6.9	1
PRJNA295431	SRR2353440	miRNA-seq	Single				1	68		7.3	1
PRJNA295431	SRR2353441	miRNA-seq	Single				1	95		7	1
PRJNA295431	SRR2353442	miRNA-seq	Single				1	74		7.5	0
PRJNA295431	SRR2353443	miRNA-seq	Single				1	68		7.9	0
PRJNA295431	SRR2353444	miRNA-seq	Single				1	79		6.9	0
PRJNA295431	SRR2353445	miRNA-seq	Single				1	70		6.7	0
PRJNA557205	SRR9854771	RNA-seq	Single				1	77		7.6	
PRJNA557205	SRR9854772	RNA-seq	Single				1	77		7.6	
PRJNA557205	SRR9854773	RNA-seq	Single				1	77		7.6	
PRJNA557205	SRR9854774	RNA-seq	Single				1	77		7.6	
PRJNA557205	SRR9854775	RNA-seq	Single				1	80		6.6	
PRJNA557205	SRR9854776	RNA-seq	Single				1	80		6.6	
PRJNA557205	SRR9854777	RNA-seq	Single				1	80		6.6	
PRJNA557205	SRR9854778	RNA-seq	Single				1	80		6.6	
PRJNA557205	SRR9854779	RNA-seq	Single				2	78		6.7	
PRJNA557205	SRR9854780	RNA-seq	Single				2	78		6.7	
PRJNA557205	SRR9854781	RNA-seq	Single				2	78		6.7	

PRJNA557205	SRR9854782	RNA-seq	Single			2	78		6.7	
PRJNA557205	SRR9854783	RNA-seq	Single			1	73		6	
PRJNA557205	SRR9854784	RNA-seq	Single			1	73		6	
PRJNA557205	SRR9854785	RNA-seq	Single			1	73		6	
PRJNA557205	SRR9854786	RNA-seq	Single			1	73		6	
PRJNA557205	SRR9854787	RNA-seq	Single			1	75		7.7	
PRJNA557205	SRR9854788	RNA-seq	Single			1	75		7.7	
PRJNA557205	SRR9854789	RNA-seq	Single			1	75		7.7	
PRJNA557205	SRR9854790	RNA-seq	Single			1	75		7.7	
PRJNA557205	SRR9854791	RNA-seq	Single			1	71		7.5	
PRJNA557205	SRR9854792	RNA-seq	Single			1	71		7.5	
PRJNA557205	SRR9854793	RNA-seq	Single			1	71		7.5	
PRJNA557205	SRR9854794	RNA-seq	Single			1	71		7.5	
PRJNA557205	SRR9854795	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854796	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854797	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854798	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854799	RNA-seq	Single			1	78		6.4	
PRJNA557205	SRR9854800	RNA-seq	Single			1	78		6.4	
PRJNA557205	SRR9854801	RNA-seq	Single			1	78		6.4	
PRJNA557205	SRR9854802	RNA-seq	Single			1	78		6.4	
PRJNA557205	SRR9854803	RNA-seq	Single			1	81		7.4	
PRJNA557205	SRR9854804	RNA-seq	Single			1	81		7.4	
PRJNA557205	SRR9854805	RNA-seq	Single			1	81		7.4	
PRJNA557205	SRR9854806	RNA-seq	Single			1	81		7.4	
PRJNA557205	SRR9854807	RNA-seq	Single			1	72		7.9	
PRJNA557205	SRR9854808	RNA-seq	Single			1	72		7.9	
PRJNA557205	SRR9854809	RNA-seq	Single			1	72		7.9	
PRJNA557205	SRR9854810	RNA-seq	Single			1	72		7.9	
PRJNA557205	SRR9854811	RNA-seq	Single			1	82		6.7	
PRJNA557205	SRR9854812	RNA-seq	Single			1	82		6.7	
PRJNA557205	SRR9854813	RNA-seq	Single			1	82		6.7	
PRJNA557205	SRR9854814	RNA-seq	Single			1	82		6.7	
PRJNA557205	SRR9854815	RNA-seq	Single			1	70		6.5	
PRJNA557205	SRR9854816	RNA-seq	Single			1	70		6.5	
PRJNA557205	SRR9854817	RNA-seq	Single			1	70		6.5	
PRJNA557205	SRR9854818	RNA-seq	Single			1	70		6.5	
PRJNA557205	SRR9854819	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854820	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854821	RNA-seq	Single			1	78		7.6	
PRJNA557205	SRR9854822	RNA-seq	Single			1	78		7.6	

PRJNA557205	SRR9854823	RNA-seq	Single			1	67		8.1	
PRJNA557205	SRR9854824	RNA-seq	Single			1	67		8.1	
PRJNA557205	SRR9854825	RNA-seq	Single			1	67		8.1	
PRJNA557205	SRR9854826	RNA-seq	Single			1	67		8.1	
PRJNA557205	SRR9854827	RNA-seq	Single			1	78		7	
PRJNA557205	SRR9854828	RNA-seq	Single			1	78		7	
PRJNA557205	SRR9854829	RNA-seq	Single			1	78		7	
PRJNA557205	SRR9854830	RNA-seq	Single			1	78		7	
PRJNA557205	SRR9854831	RNA-seq	Single			1	69		6.2	
PRJNA557205	SRR9854832	RNA-seq	Single			1	69		6.2	
PRJNA557205	SRR9854833	RNA-seq	Single			1	69		6.2	
PRJNA557205	SRR9854834	RNA-seq	Single			1	69		6.2	
PRJNA557205	SRR9854835	RNA-seq	Single			2	83		6.9	
PRJNA557205	SRR9854836	RNA-seq	Single			2	83		6.9	
PRJNA557205	SRR9854837	RNA-seq	Single			2	83		6.9	
PRJNA557205	SRR9854838	RNA-seq	Single			2	83		6.9	
PRJNA557205	SRR9854839	RNA-seq	Single			1	81		7.2	
PRJNA557205	SRR9854840	RNA-seq	Single			1	81		7.2	
PRJNA557205	SRR9854841	RNA-seq	Single			1	81		7.2	
PRJNA557205	SRR9854842	RNA-seq	Single			1	81		7.2	
PRJNA557205	SRR9854843	RNA-seq	Single			1	74		6.9	
PRJNA557205	SRR9854844	RNA-seq	Single			1	74		6.9	
PRJNA557205	SRR9854845	RNA-seq	Single			1	74		6.9	
PRJNA557205	SRR9854846	RNA-seq	Single			1	74		6.9	
PRJNA557205	SRR9854847	RNA-seq	Single			2	66		8.2	
PRJNA557205	SRR9854848	RNA-seq	Single			2	66		8.2	
PRJNA557205	SRR9854849	RNA-seq	Single			2	66		8.2	
PRJNA557205	SRR9854850	RNA-seq	Single			2	66		8.2	
PRJNA557205	SRR9854851	RNA-seq	Single			1	73		6.7	
PRJNA557205	SRR9854852	RNA-seq	Single			1	73		6.7	
PRJNA557205	SRR9854853	RNA-seq	Single			1	73		6.7	
PRJNA557205	SRR9854854	RNA-seq	Single			1	73		6.7	
PRJNA557205	SRR9854855	RNA-seq	Single			1	81		7.9	
PRJNA557205	SRR9854856	RNA-seq	Single			1	81		7.9	
PRJNA557205	SRR9854857	RNA-seq	Single			1	81		7.9	
PRJNA557205	SRR9854858	RNA-seq	Single			1	81		7.9	
PRJNA557205	SRR9854859	RNA-seq	Single			1	72		7.7	
PRJNA557205	SRR9854860	RNA-seq	Single			1	72		7.7	
PRJNA557205	SRR9854861	RNA-seq	Single			1	72		7.7	
PRJNA557205	SRR9854862	RNA-seq	Single			1	72		7.7	
PRJNA557205	SRR9854863	RNA-seq	Single			1	77		6.8	

PRJNA557205	SRR9854864	RNA-seq	Single			1	77		6.8	
PRJNA557205	SRR9854865	RNA-seq	Single			1	77		6.8	
PRJNA557205	SRR9854866	RNA-seq	Single			1	77		6.8	
PRJEB5212	ERR409878	RNA-seq	Single	Ctrl	DLPFC/BA9	1	43			
PRJEB5212	ERR409879	RNA-seq	Single			1	80			
PRJEB5212	ERR409882	RNA-seq	Single			1	53			
PRJEB5212	ERR409883	RNA-seq	Single			1	46			
PRJEB5212	ERR409887	RNA-seq	Single			1	76			
PRJEB5212	ERR409888	RNA-seq	Single			1	57			
PRJEB5212	ERR409891	RNA-seq	Single			1	79			
PRJEB5212	ERR409893	RNA-seq	Single			1	40			
PRJEB5212	ERR409896	RNA-seq	Single			1	57			
PRJEB5212	ERR409898	RNA-seq	Single			1	93			
PRJEB5212	ERR409900	RNA-seq	Single			1	44			
PRJNA143627	SRR306840	RNA-seq	Single			1			8.6	
PRJNA143627	SRR306841	RNA-seq	Single			1			8.5	
PRJNA143627	SRR306842	RNA-seq	Single			1			7.8	
PRJNA185476	SRR646457	RNA-seq	Single			2				
PRJNA185476	SRR646458	RNA-seq	Single			2				
PRJNA185476	SRR646459	RNA-seq	Single			2				
PRJNA185476	SRR646460	RNA-seq	Single			2				
PRJNA283498	SRR2015713	RNA-seq	Paired			1	73	3	7.7	
PRJNA283498	SRR2015714	RNA-seq	Paired			1	91	2	7.9	
PRJNA283498	SRR2015715	RNA-seq	Paired			1	82	3	8.6	
PRJNA283498	SRR2015716	RNA-seq	Paired			1	97	3	9.1	
PRJNA283498	SRR2015717	RNA-seq	Paired			1	86	1	8.6	
PRJNA283498	SRR2015718	RNA-seq	Paired			1	91	4	8.7	
PRJNA283498	SRR2015719	RNA-seq	Paired			1	81	1	6	
PRJNA283498	SRR2015720	RNA-seq	Paired			1	79	1	8.4	
PRJNA283498	SRR2015721	RNA-seq	Paired			1	63	2	6.5	
PRJNA283498	SRR2015722	RNA-seq	Paired			1	66		7.1	
PRJNA283498	SRR2015723	RNA-seq	Paired			1	69		7.8	
PRJNA283498	SRR2015724	RNA-seq	Paired			1	79		8	
PRJNA283498	SRR2015725	RNA-seq	Paired			1	61	1	8.2	
PRJNA283498	SRR2015726	RNA-seq	Paired			1	58	0	8.4	
PRJNA283498	SRR2015727	RNA-seq	Paired			1	70	1	8.2	
PRJNA283498	SRR2015728	RNA-seq	Paired			1	66	1	8.5	
PRJNA283498	SRR2015729	RNA-seq	Paired			1	73	1	7.5	
PRJNA283498	SRR2015730	RNA-seq	Paired			1	60		7.9	
PRJNA283498	SRR2015731	RNA-seq	Paired			1	76	1	7.3	
PRJNA283498	SRR2015732	RNA-seq	Paired			1	61	0	7.8	

PRJNA283498	SRR2015733	RNA-seq	Paired			1	62	1	6.6	
PRJNA283498	SRR2015734	RNA-seq	Paired			1	69	1	8.7	
PRJNA283498	SRR2015735	RNA-seq	Paired			1	61	1	8.1	
PRJNA283498	SRR2015736	RNA-seq	Paired			1	88	2	7.1	
PRJNA283498	SRR2015737	RNA-seq	Paired			1	93		6.4	
PRJNA283498	SRR2015738	RNA-seq	Paired			1	53	0	7.3	
PRJNA283498	SRR2015739	RNA-seq	Paired			1	57	1	8.3	
PRJNA283498	SRR2015740	RNA-seq	Paired			1	46	0	7.6	
PRJNA283498	SRR2015741	RNA-seq	Paired			1	57	1	7.7	
PRJNA283498	SRR2015742	RNA-seq	Paired			1	80	1	7.3	
PRJNA283498	SRR2015743	RNA-seq	Paired			1	74	2	8.5	
PRJNA283498	SRR2015744	RNA-seq	Paired			1	69	1	8.4	
PRJNA283498	SRR2015745	RNA-seq	Paired			1	76	1	7.5	
PRJNA283498	SRR2015746	RNA-seq	Paired			1	78	2	7.6	
PRJNA283498	SRR2015747	RNA-seq	Paired			1	87	4	8.7	
PRJNA283498	SRR2015748	RNA-seq	Paired			1	86	2	8.7	
PRJNA283498	SRR2015749	RNA-seq	Paired			1	54		8.3	
PRJNA283498	SRR2015750	RNA-seq	Paired			1	68	0	6.3	
PRJNA283498	SRR2015751	RNA-seq	Paired			1	52	0	7.4	
PRJNA283498	SRR2015752	RNA-seq	Paired			1	46	0	8.2	
PRJNA283498	SRR2015753	RNA-seq	Paired			1	55		7.6	
PRJNA283498	SRR2015754	RNA-seq	Paired			1	57	1	7.8	
PRJNA283498	SRR2015755	RNA-seq	Paired			1	66	1	8.4	
PRJNA283498	SRR2015756	RNA-seq	Paired			1	64	1	8.7	
PRJNA290524	SRR2121907	RNA-seq	Paired							
PRJNA557205	SRR9854867	RNA-seq	Single			2	79		7.5	
PRJNA557205	SRR9854868	RNA-seq	Single			2	79		7.5	
PRJNA557205	SRR9854869	RNA-seq	Single			2	79		7.5	
PRJNA557205	SRR9854870	RNA-seq	Single			2	79		7.5	
PRJNA557205	SRR9854871	RNA-seq	Single			1	82		8.3	
PRJNA557205	SRR9854872	RNA-seq	Single			1	82		8.3	
PRJNA557205	SRR9854873	RNA-seq	Single			1	82		8.3	
PRJNA557205	SRR9854874	RNA-seq	Single			1	82		8.3	
PRJNA557205	SRR9854875	RNA-seq	Single			1	70		7.7	
PRJNA557205	SRR9854876	RNA-seq	Single			1	70		7.7	
PRJNA557205	SRR9854877	RNA-seq	Single			1	70		7.7	
PRJNA557205	SRR9854878	RNA-seq	Single			1	70		7.7	
PRJNA557205	SRR9854879	RNA-seq	Single			1	77		7.1	
PRJNA557205	SRR9854880	RNA-seq	Single			1	77		7.1	
PRJNA557205	SRR9854881	RNA-seq	Single			1	77		7.1	
PRJNA557205	SRR9854882	RNA-seq	Single			1	77		7.1	

PRJNA557205	SRR9854883	RNA-seq	Single			1	73		7.8	
PRJNA557205	SRR9854884	RNA-seq	Single			1	73		7.8	
PRJNA557205	SRR9854885	RNA-seq	Single			1	73		7.8	
PRJNA557205	SRR9854886	RNA-seq	Single			1	73		7.8	
PRJNA557205	SRR9854887	RNA-seq	Single			1	73		7.1	
PRJNA557205	SRR9854888	RNA-seq	Single			1	73		7.1	
PRJNA557205	SRR9854889	RNA-seq	Single			1	73		7.1	
PRJNA557205	SRR9854890	RNA-seq	Single			1	73		7.1	
PRJNA557205	SRR9854891	RNA-seq	Single			1	68		8.1	
PRJNA557205	SRR9854892	RNA-seq	Single			1	68		8.1	
PRJNA557205	SRR9854893	RNA-seq	Single			1	68		8.1	
PRJNA557205	SRR9854894	RNA-seq	Single			1	68		8.1	
PRJNA557205	SRR9854895	RNA-seq	Single			1	68		7.6	
PRJNA557205	SRR9854896	RNA-seq	Single			1	68		7.6	
PRJNA557205	SRR9854897	RNA-seq	Single			1	68		7.6	
PRJNA557205	SRR9854898	RNA-seq	Single			1	68		7.6	
PRJNA557205	SRR9854899	RNA-seq	Single			1	84		8.2	
PRJNA557205	SRR9854900	RNA-seq	Single			1	84		8.2	
PRJNA557205	SRR9854901	RNA-seq	Single			1	84		8.2	
PRJNA557205	SRR9854902	RNA-seq	Single			1	84		8.2	
PRJNA557205	SRR9854903	RNA-seq	Single			1	71		8.8	
PRJNA557205	SRR9854904	RNA-seq	Single			1	71		8.8	
PRJNA557205	SRR9854905	RNA-seq	Single			1	71		8.8	
PRJNA557205	SRR9854906	RNA-seq	Single			1	71		8.8	
PRJNA557205	SRR9854907	RNA-seq	Single			1	83		7.3	
PRJNA557205	SRR9854908	RNA-seq	Single			1	83		7.3	
PRJNA557205	SRR9854909	RNA-seq	Single			1	83		7.3	
PRJNA557205	SRR9854910	RNA-seq	Single			1	83		7.3	
PRJNA557205	SRR9854911	RNA-seq	Single			1	70		8	
PRJNA557205	SRR9854912	RNA-seq	Single			1	70		8	
PRJNA557205	SRR9854913	RNA-seq	Single			1	70		8	
PRJNA557205	SRR9854914	RNA-seq	Single			1	70		8	
PRJNA505258	SRR8181157	RNA-seq	Single	AD	PBMC	1	77			
PRJNA505258	SRR8181160	RNA-seq	Single			2	75			
PRJNA505258	SRR8181164	RNA-seq	Single			1	80			
PRJNA505258	SRR8181168	RNA-seq	Single			1	79			
PRJNA505258	SRR8181172	RNA-seq	Single			2	77			
PRJNA505258	SRR8181176	RNA-seq	Single			2	76			
PRJNA505258	SRR8181180	RNA-seq	Single			1	80			
PRJNA505258	SRR8181183	RNA-seq	Single			1	74			
PRJNA505258	SRR8181185	RNA-seq	Single			1	78			

PRJNA505258	SRR8181186	RNA-seq	Single			2	72			
PRJNA505258	SRR8181192	RNA-seq	Single			1	73			
PRJNA667860	SRR12786082	AMPLICON	Paired			2	66			
PRJNA667860	SRR12786083	AMPLICON	Paired			1	76			
PRJNA667860	SRR12786084	AMPLICON	Paired			2	67			
PRJNA667860	SRR12786085	AMPLICON	Paired			2	67			
PRJNA667860	SRR12786086	AMPLICON	Paired			2	69			
PRJNA667860	SRR12786087	AMPLICON	Paired			2	68			
PRJNA667860	SRR12786088	AMPLICON	Paired			2	68			
PRJNA667860	SRR12786089	AMPLICON	Paired			2	81			
PRJNA667860	SRR12786094	AMPLICON	Paired			1	63			
PRJNA667860	SRR12786095	AMPLICON	Paired			2	78			
PRJNA667860	SRR12786096	AMPLICON	Paired			2	79			
PRJNA667860	SRR12786097	AMPLICON	Paired			1	61			
PRJNA678865	SRR13071573	RNA-seq	Paired			2	70			
PRJNA678865	SRR13071574	RNA-seq	Paired			2	73			
PRJNA678865	SRR13071575	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071576	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071577	RNA-seq	Paired			2	75			
PRJNA678865	SRR13071585	RNA-seq	Paired			2	70			
PRJNA678865	SRR13071586	RNA-seq	Paired			2	73			
PRJNA678865	SRR13071587	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071588	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071589	RNA-seq	Paired			2	75			
PRJNA505258	SRR8181194	RNA-seq	Single	Ctrl	PBMC	2	71			
PRJNA505258	SRR8181195	RNA-seq	Single			2	75			
PRJNA505258	SRR8181196	RNA-seq	Single			2	79			
PRJNA505258	SRR8181203	RNA-seq	Single			1	76			
PRJNA505258	SRR8181207	RNA-seq	Single			1	73			
PRJNA505258	SRR8181210	RNA-seq	Single			2	71			
PRJNA505258	SRR8181214	RNA-seq	Single			1	72			
PRJNA505258	SRR8181219	RNA-seq	Single			1	77			
PRJNA505258	SRR8181220	RNA-seq	Single			2	74			
PRJNA505258	SRR8181221	RNA-seq	Single			2	78			
PRJNA505258	SRR8181222	RNA-seq	Single			1	71			
PRJNA667860	SRR12786090	AMPLICON	Paired			1	72			
PRJNA667860	SRR12786091	AMPLICON	Paired			1	78			
PRJNA667860	SRR12786092	AMPLICON	Paired			1	78			
PRJNA667860	SRR12786093	AMPLICON	Paired			1	76			
PRJNA678865	SRR13071578	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071579	RNA-seq	Paired			2	76			

PRJNA678865	SRR13071584	RNA-seq	Paired			1	74			
PRJNA678865	SRR13071590	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071591	RNA-seq	Paired			2	76			
PRJNA678865	SRR13071592	RNA-seq	Paired			1	74			
PRJNA678865	SRR13071593	RNA-seq	Paired			1	61			
PRJNA678865	SRR13071594	RNA-seq	Paired			1	58			
PRJNA678865	SRR13071595	RNA-seq	Paired			1	61			
PRJNA678865	SRR13071596	RNA-seq	Paired			1	58			
PRJNA556869	SRR9844387	ncRNA-seq	Single	PD	PBMC	2	64			
PRJNA556869	SRR9844388	ncRNA-seq	Single			1	65			
PRJNA556869	SRR9844389	ncRNA-seq	Single			2	69			
PRJNA556869	SRR9844390	ncRNA-seq	Single			1	70			
PRJNA556869	SRR9844391	ncRNA-seq	Single			1	81			
PRJNA556869	SRR9844392	ncRNA-seq	Single			2	55			
PRJNA556869	SRR9844393	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844394	ncRNA-seq	Single			1	70			
PRJNA556869	SRR9844395	ncRNA-seq	Single			1	58			
PRJNA556869	SRR9844396	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844397	ncRNA-seq	Single			1	58			
PRJNA556869	SRR9844398	ncRNA-seq	Single			2	70			
PRJNA556869	SRR9844399	ncRNA-seq	Single			1	60			
PRJNA556869	SRR9844400	ncRNA-seq	Single			2	68			
PRJNA556869	SRR9844401	ncRNA-seq	Single			1	70			
PRJNA556869	SRR9844402	ncRNA-seq	Single			1	69			
PRJNA556869	SRR9844403	ncRNA-seq	Single			2	60			
PRJNA556869	SRR9844404	ncRNA-seq	Single			2	60			
PRJNA556869	SRR9844405	ncRNA-seq	Single			2	72			
PRJNA556869	SRR9844406	ncRNA-seq	Single			1	64			
PRJNA556869	SRR9844407	ncRNA-seq	Single			2	84			
PRJNA556869	SRR9844408	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844409	ncRNA-seq	Single			1	57			
PRJNA556869	SRR9844410	ncRNA-seq	Single			2	65			
PRJNA556869	SRR9844411	ncRNA-seq	Single			1	74			
PRJNA556869	SRR9844412	ncRNA-seq	Single			1	52			
PRJNA556869	SRR9844413	ncRNA-seq	Single			2	42			
PRJNA556869	SRR9844414	ncRNA-seq	Single			1	60			
PRJNA556869	SRR9844415	ncRNA-seq	Single			1	78			
PRJNA556869	SRR9844416	ncRNA-seq	Single			1	58			
PRJNA556869	SRR9844417	ncRNA-seq	Single			1	62			
PRJNA556869	SRR9844418	ncRNA-seq	Single			1	70			
PRJNA556869	SRR9844419	ncRNA-seq	Single			1	56			

PRJNA556869	SRR9844420	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844421	ncRNA-seq	Single			1	73			
PRJNA556869	SRR9844422	ncRNA-seq	Single			1	65			
PRJNA556869	SRR9844423	ncRNA-seq	Single			1	68			
PRJNA556869	SRR9844424	ncRNA-seq	Single			1	77			
PRJNA556869	SRR9844425	ncRNA-seq	Single			1	55			
PRJNA556869	SRR9844426	ncRNA-seq	Single			1	64			
PRJNA556869	SRR9844427	ncRNA-seq	Single			2	76			
PRJNA556869	SRR9844428	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844429	ncRNA-seq	Single			1	69			
PRJNA556869	SRR9844430	ncRNA-seq	Single			2	58			
PRJNA556869	SRR9844431	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844432	ncRNA-seq	Single			2	71			
PRJNA556869	SRR9844433	ncRNA-seq	Single			1	59			
PRJNA556869	SRR9844434	ncRNA-seq	Single	Ctrl	PBMC	2	64			
PRJNA556869	SRR9844435	ncRNA-seq	Single			2	62			
PRJNA556869	SRR9844436	ncRNA-seq	Single			1	72			
PRJNA556869	SRR9844437	ncRNA-seq	Single			1	58			
PRJNA556869	SRR9844438	ncRNA-seq	Single			1	64			
PRJNA556869	SRR9844439	ncRNA-seq	Single			2	62			
PRJNA556869	SRR9844440	ncRNA-seq	Single			1	63			
PRJNA556869	SRR9844441	ncRNA-seq	Single			2	70			
PRJNA556869	SRR9844442	ncRNA-seq	Single			2	57			
PRJNA556869	SRR9844443	ncRNA-seq	Single			2	54			
PRJNA556869	SRR9844444	ncRNA-seq	Single			1	56			
PRJNA556869	SRR9844445	ncRNA-seq	Single			1	50			
PRJNA556869	SRR9844446	ncRNA-seq	Single			2	61			
PRJNA556869	SRR9844447	ncRNA-seq	Single			2	59			
PRJNA556869	SRR9844448	ncRNA-seq	Single			1	62			
PRJNA556869	SRR9844449	ncRNA-seq	Single			1	59			
PRJNA556869	SRR9844450	ncRNA-seq	Single			1	78			
PRJNA556869	SRR9844451	ncRNA-seq	Single			2	60			
PRJNA556869	SRR9844452	ncRNA-seq	Single			2	71			
PRJNA556869	SRR9844453	ncRNA-seq	Single			2	51			
PRJNA556869	SRR9844454	ncRNA-seq	Single			2	56			
PRJNA556869	SRR9844455	ncRNA-seq	Single			2	60			
PRJNA556869	SRR9844456	ncRNA-seq	Single			2	65			
PRJNA556869	SRR9844457	ncRNA-seq	Single			2	73			
PRJNA556869	SRR9844458	ncRNA-seq	Single			1	76			
PRJNA556869	SRR9844459	ncRNA-seq	Single			1	73			
PRJNA556869	SRR9844460	ncRNA-seq	Single			1	68			

PRJNA556869	SRR9844361	ncRNA-seq	Single			1	72			
PRJNA556869	SRR9844362	ncRNA-seq	Single			2	53			
PRJNA556869	SRR9844363	ncRNA-seq	Single			1	59			
PRJNA556869	SRR9844364	ncRNA-seq	Single			1	69			
PRJNA556869	SRR9844365	ncRNA-seq	Single			2	62			
PRJNA556869	SRR9844366	ncRNA-seq	Single			1	84			
PRJNA556869	SRR9844367	ncRNA-seq	Single			1	75			
PRJNA556869	SRR9844368	ncRNA-seq	Single			1	63			
PRJNA556869	SRR9844369	ncRNA-seq	Single			1	63			
PRJNA556869	SRR9844370	ncRNA-seq	Single			1	63			
PRJNA556869	SRR9844371	ncRNA-seq	Single			2	70			
PRJNA556869	SRR9844372	ncRNA-seq	Single			2	64			
PRJNA556869	SRR9844373	ncRNA-seq	Single			2	60			
PRJNA556869	SRR9844374	ncRNA-seq	Single			1	60			
PRJNA556869	SRR9844375	ncRNA-seq	Single			1	65			
PRJNA556869	SRR9844376	ncRNA-seq	Single			1	64			
PRJNA556869	SRR9844377	ncRNA-seq	Single			2	74			
PRJNA556869	SRR9844378	ncRNA-seq	Single			1	68			
PRJNA556869	SRR9844379	ncRNA-seq	Single			2	66			
PRJNA556869	SRR9844380	ncRNA-seq	Single			1	56			
PRJNA556869	SRR9844381	ncRNA-seq	Single			2	64			
PRJNA556869	SRR9844382	ncRNA-seq	Single			2	68			
PRJNA556869	SRR9844383	ncRNA-seq	Single			1	70			
PRJNA556869	SRR9844384	ncRNA-seq	Single			1	68			
PRJNA556869	SRR9844385	ncRNA-seq	Single			1	69			
PRJNA556869	SRR9844386	ncRNA-seq	Single			1	69			
PRJNA201039	SRR837437	RNA-seq	Single	AD	WB	2	77			
PRJNA201039	SRR837438	RNA-seq	Single			1	74			
PRJNA201039	SRR837439	RNA-seq	Single			1	68			
PRJNA201039	SRR837440	RNA-seq	Single			2	75			
PRJNA201039	SRR837441	RNA-seq	Single			1	74			
PRJNA201039	SRR837442	RNA-seq	Single			1	76			
PRJNA201039	SRR837443	RNA-seq	Single			1	79			
PRJNA201039	SRR837444	RNA-seq	Single			2	75			
PRJNA201039	SRR837445	RNA-seq	Single			1	77			
PRJNA201039	SRR837446	RNA-seq	Single			2	72			
PRJNA201039	SRR837447	RNA-seq	Single			1	76			
PRJNA201039	SRR837448	RNA-seq	Single			1	61			
PRJNA201039	SRR837449	RNA-seq	Single			2	51			
PRJNA201039	SRR837450	RNA-seq	Single			2	55			
PRJNA201039	SRR837451	RNA-seq	Single			2	70			

PRJNA201039	SRR837452	RNA-seq	Single			2	78			
PRJNA201039	SRR837453	RNA-seq	Single			1	81			
PRJNA201039	SRR837454	RNA-seq	Single			2	72			
PRJNA201039	SRR837455	RNA-seq	Single			1	57			
PRJNA201039	SRR837456	RNA-seq	Single			1	75			
PRJNA201039	SRR837457	RNA-seq	Single			2	70			
PRJNA201039	SRR837458	RNA-seq	Single			2	75			
PRJNA201039	SRR837459	RNA-seq	Single			2	69			
PRJNA201039	SRR837460	RNA-seq	Single			2	68			
PRJNA201039	SRR837461	RNA-seq	Single			2	71			
PRJNA201039	SRR837462	RNA-seq	Single			1	74			
PRJNA201039	SRR837463	RNA-seq	Single			1	72			
PRJNA201039	SRR837464	RNA-seq	Single			2	69			
PRJNA201039	SRR837465	RNA-seq	Single			2	63			
PRJNA201039	SRR837466	RNA-seq	Single			1	80			
PRJNA201039	SRR837467	RNA-seq	Single			1	62			
PRJNA201039	SRR837468	RNA-seq	Single			1	75			
PRJNA201039	SRR837469	RNA-seq	Single			2	56			
PRJNA201039	SRR837470	RNA-seq	Single			2	57			
PRJNA201039	SRR837493	RNA-seq	Single			1	80			
PRJNA201039	SRR837494	RNA-seq	Single			1	81			
PRJNA201039	SRR837495	RNA-seq	Single			2	76			
PRJNA201039	SRR837496	RNA-seq	Single			2	60			
PRJNA201039	SRR837497	RNA-seq	Single			2	63			
PRJNA201039	SRR837498	RNA-seq	Single			2	64			
PRJNA201039	SRR837499	RNA-seq	Single			1	65			
PRJNA201039	SRR837500	RNA-seq	Single			2	74			
PRJNA201039	SRR837501	RNA-seq	Single			2	59			
PRJNA201039	SRR837502	RNA-seq	Single			2	72			
PRJNA201039	SRR837503	RNA-seq	Single			1	63			
PRJNA201039	SRR837504	RNA-seq	Single			1	78			
PRJNA201039	SRR837505	RNA-seq	Single			1	78			
PRJNA201039	SRR837506	RNA-seq	Single			1	79			
PRJNA675864	SRR13019404	RNA-seq	Paired							
PRJNA675864	SRR13019405	RNA-seq	Paired							
PRJNA675864	SRR13019406	RNA-seq	Paired							
PRJNA675864	SRR13019407	RNA-seq	Paired							
PRJNA675864	SRR13019408	RNA-seq	Paired							
PRJNA675864	SRR13019409	RNA-seq	Paired							
PRJNA201039	SRR837471	RNA-seq	Single	Ctrl	WB	1	76			
PRJNA201039	SRR837472	RNA-seq	Single			2	60			

PRJNA201039	SRR837473	RNA-seq	Single	PD	WB	2	60			
PRJNA201039	SRR837474	RNA-seq	Single			2	76			
PRJNA201039	SRR837475	RNA-seq	Single			2	60			
PRJNA201039	SRR837476	RNA-seq	Single			1	83			
PRJNA201039	SRR837477	RNA-seq	Single			1	61			
PRJNA201039	SRR837478	RNA-seq	Single			1	77			
PRJNA201039	SRR837479	RNA-seq	Single			1	63			
PRJNA201039	SRR837480	RNA-seq	Single			1	75			
PRJNA201039	SRR837481	RNA-seq	Single			1	63			
PRJNA201039	SRR837482	RNA-seq	Single			1	64			
PRJNA201039	SRR837483	RNA-seq	Single			2	71			
PRJNA201039	SRR837484	RNA-seq	Single			2	60			
PRJNA201039	SRR837485	RNA-seq	Single			1	75			
PRJNA201039	SRR837486	RNA-seq	Single			2	60			
PRJNA201039	SRR837487	RNA-seq	Single			2	66			
PRJNA201039	SRR837488	RNA-seq	Single			2	60			
PRJNA201039	SRR837489	RNA-seq	Single			1	61			
PRJNA201039	SRR837490	RNA-seq	Single			2	74			
PRJNA201039	SRR837491	RNA-seq	Single			2	71			
PRJNA201039	SRR837492	RNA-seq	Single			1	60			
PRJNA675864	SRR13019410	RNA-seq	Paired							
PRJNA675864	SRR13019411	RNA-seq	Paired							
PRJNA675864	SRR13019412	RNA-seq	Paired							
PRJNA675864	SRR13019413	RNA-seq	Paired							
PRJNA675864	SRR13019414	RNA-seq	Paired							
PRJNA675864	SRR13019415	RNA-seq	Paired							
PRJNA675864	SRR13019416	RNA-seq	Paired							
PRJEB8960	ERR833219	RNA-seq	Single	PD	WB	2	59			
PRJEB8960	ERR833220	RNA-seq	Single			2	59			
PRJEB8960	ERR833221	RNA-seq	Single			2	59			
PRJEB8960	ERR833222	RNA-seq	Single			2	59			
PRJEB8960	ERR833227	RNA-seq	Single			1	53			
PRJEB8960	ERR833228	RNA-seq	Single			1	53			
PRJEB8960	ERR833229	RNA-seq	Single			1	53			
PRJEB8960	ERR833230	RNA-seq	Single			1	53			
PRJEB8960	ERR833239	RNA-seq	Single			1	50			
PRJEB8960	ERR833240	RNA-seq	Single			1	50			
PRJEB8960	ERR833241	RNA-seq	Single			1	50			
PRJEB8960	ERR833242	RNA-seq	Single			1	50			
PRJEB8960	ERR833251	RNA-seq	Single			1	47			
PRJEB8960	ERR833252	RNA-seq	Single			1	47			

PRJEB8960	ERR833253	RNA-seq	Single			1	47			
PRJEB8960	ERR833254	RNA-seq	Single			1	47			
PRJEB8960	ERR833255	RNA-seq	Single			1	50			
PRJEB8960	ERR833256	RNA-seq	Single			1	50			
PRJEB8960	ERR833257	RNA-seq	Single			1	50			
PRJEB8960	ERR833258	RNA-seq	Single			1	50			
PRJNA675864	SRR13019398	RNA-seq	Paired			1	62			
PRJNA675864	SRR13019399	RNA-seq	Paired			1	68			
PRJNA675864	SRR13019400	RNA-seq	Paired			1	56			
PRJNA675864	SRR13019401	RNA-seq	Paired			1	67			
PRJNA675864	SRR13019402	RNA-seq	Paired			2	64			
PRJNA675864	SRR13019403	RNA-seq	Paired			2	58			
PRJNA693202	SRR13482599	RNA-seq	Paired							
PRJNA693202	SRR13482600	RNA-seq	Paired							
PRJNA693202	SRR13482601	RNA-seq	Paired							
PRJNA693202	SRR13482602	RNA-seq	Paired							
PRJNA693202	SRR13482603	RNA-seq	Paired							
PRJNA693202	SRR13482604	RNA-seq	Paired							
PRJNA693202	SRR13482605	RNA-seq	Paired							
PRJNA693202	SRR13482606	RNA-seq	Paired							
PRJNA693202	SRR13482607	RNA-seq	Paired							
PRJNA693202	SRR13482608	RNA-seq	Paired							
PRJNA693202	SRR13482609	RNA-seq	Paired							
PRJNA693202	SRR13482610	RNA-seq	Paired							
PRJEB8960	ERR833223	RNA-seq	Single	Ctrl	WB	2	54			
PRJEB8960	ERR833224	RNA-seq	Single			2	54			
PRJEB8960	ERR833225	RNA-seq	Single			2	54			
PRJEB8960	ERR833226	RNA-seq	Single			2	54			
PRJEB8960	ERR833231	RNA-seq	Single			1	45			
PRJEB8960	ERR833232	RNA-seq	Single			1	45			
PRJEB8960	ERR833233	RNA-seq	Single			1	45			
PRJEB8960	ERR833234	RNA-seq	Single			1	45			
PRJEB8960	ERR833235	RNA-seq	Single			1	57			
PRJEB8960	ERR833236	RNA-seq	Single			1	57			
PRJEB8960	ERR833237	RNA-seq	Single			1	57			
PRJEB8960	ERR833238	RNA-seq	Single			1	57			
PRJEB8960	ERR833243	RNA-seq	Single			2	42			
PRJEB8960	ERR833244	RNA-seq	Single			2	42			
PRJEB8960	ERR833245	RNA-seq	Single			2	42			
PRJEB8960	ERR833246	RNA-seq	Single			2	42			
PRJEB8960	ERR833247	RNA-seq	Single			1	42			

PRJEB8960	ERR833248	RNA-seq	Single			1	42			
PRJEB8960	ERR833249	RNA-seq	Single			1	42			
PRJEB8960	ERR833250	RNA-seq	Single			1	42			
PRJNA675864	SRR13019394	RNA-seq	Paired			53	2			
PRJNA675864	SRR13019395	RNA-seq	Paired			61	1			
PRJNA675864	SRR13019396	RNA-seq	Paired			45	1			
PRJNA675864	SRR13019397	RNA-seq	Paired			56	1			
PRJNA675864	SRR13019410	RNA-seq	Paired			60	2			
PRJNA675864	SRR13019411	RNA-seq	Paired			57	1			
PRJNA675864	SRR13019412	RNA-seq	Paired			48	2			
PRJNA675864	SRR13019413	RNA-seq	Paired			59	1			
PRJNA675864	SRR13019414	RNA-seq	Paired			51	1			
PRJNA675864	SRR13019415	RNA-seq	Paired			44	1			
PRJNA675864	SRR13019416	RNA-seq	Paired			50	1			
PRJNA693202	SRR13482611	RNA-seq	Paired							
PRJNA693202	SRR13482612	RNA-seq	Paired							
PRJNA693202	SRR13482613	RNA-seq	Paired							
PRJNA693202	SRR13482614	RNA-seq	Paired							
PRJNA693202	SRR13482615	RNA-seq	Paired							
PRJNA693202	SRR13482616	RNA-seq	Paired							
PRJNA693202	SRR13482617	RNA-seq	Paired							
PRJNA693202	SRR13482618	RNA-seq	Paired							
PRJNA693202	SRR13482619	RNA-seq	Paired							
PRJNA693202	SRR13482620	RNA-seq	Paired							
PRJNA693202	SRR13482621	RNA-seq	Paired							
PRJNA693202	SRR13482622	RNA-seq	Paired							
PRJNA693202	SRR13482623	RNA-seq	Paired							
PRJNA693202	SRR13482624	RNA-seq	Paired							

Appendix 2 Table provides a description of the selected SRA samples through samples subtyping analysis.

The SRA publicly accessioned objects available in the meta-data: Project (study): describes the goal of the research effort. SID (SRR#): a manifest of data file(s) that are derived from sequencing a library. Strategy: describes the sequencing strategy. Layout: implies the paired-end and single-end RNA-sequencing per each sample.

Type of tissue and the related human disease can be found in the Tissue and Condition column, respectively. Related physiological traits can be found in Age, Sex, Braak, RIN, and Dementia column.

Project	SID	Strategy	Layout	Condition	Tissue	Age	Sex	Braak	RIN	Dementia
PRJNA232671	SRR1063418	RNA-seq	Single	AD	DLPFC/BA9	100	2	4		
PRJNA232671	SRR1063419	RNA-seq	Single			88	1	5		
PRJNA232671	SRR1063420	RNA-seq	Single			85	1	5		
PRJNA210613	SRR1103943	miRNA-seq	Single			79	1	5		
PRJNA210613	SRR1103944	miRNA-seq	Single			80	2	6		
PRJNA210613	SRR1103945	miRNA-seq	Single			67	1	6		
PRJNA210613	SRR1103946	miRNA-seq	Single			67	2	7		
PRJNA210613	SRR1103947	miRNA-seq	Single			68	2	8		
PRJNA210613	SRR1103948	miRNA-seq	Single			72	2	9		
PRJNA434441	SRR6742933	RNA-seq	Single			88	1	5		
PRJNA434441	SRR6742934	RNA-seq	Single			88	1	5		
PRJNA434441	SRR6742937	RNA-seq	Single			85	1	6		
PRJNA434441	SRR6742938	RNA-seq	Single			85	1	6		
PRJNA434441	SRR6742967	RNA-seq	Single			73	2	6		
PRJNA434441	SRR6742968	RNA-seq	Single			73	2	6		
PRJNA434441	SRR6742973	RNA-seq	Single			96	2	6		
PRJNA434441	SRR6742974	RNA-seq	Single			96	2	6		
PRJNA434441	SRR6742975	RNA-seq	Single			85	2	6		
PRJNA434441	SRR6742976	RNA-seq	Single			85	2	6		
PRJNA434441	SRR6742981	RNA-seq	Single			89	2	6		
PRJNA434441	SRR6742982	RNA-seq	Single			89	2	6		
PRJNA434441	SRR6742941	RNA-seq	Single			98	1	4		
PRJNA434441	SRR6742942	RNA-seq	Single			98	1	4		
PRJNA434441	SRR6742943	RNA-seq	Single			92	1	3		
PRJNA434441	SRR6742944	RNA-seq	Single			92	1	3		
PRJNA434441	SRR6742951	RNA-seq	Single			93	1	4		
PRJNA434441	SRR6742952	RNA-seq	Single			93	1	4		
PRJNA434441	SRR6742955	RNA-seq	Single			86	1	4		
PRJNA434441	SRR6742956	RNA-seq	Single			86	1	4		

PRJNA434441	SRR6742959	RNA-seq	Single			88	2	3		
PRJNA434441	SRR6742960	RNA-seq	Single			88	2	3		
PRJNA434441	SRR6742961	RNA-seq	Single			78	1	3		
PRJNA434441	SRR6742962	RNA-seq	Single			78	1	3		
PRJNA434441	SRR6742965	RNA-seq	Single			90	2	4		
PRJNA434441	SRR6742966	RNA-seq	Single			90	2	4		
PRJNA434441	SRR6742969	RNA-seq	Single			75	2	3		
PRJNA434441	SRR6742970	RNA-seq	Single			75	2	3		
PRJNA434441	SRR6742979	RNA-seq	Single			86	2	3		
PRJNA434441	SRR6742980	RNA-seq	Single			86	2	3		
PRJNA641912	SRR12092014	ncRNA-seq	Single			84	2	5		
PRJNA641912	SRR12092015	ncRNA-seq	Single			78	1	6		
PRJNA641912	SRR12092016	ncRNA-seq	Single			98	2	5		
PRJNA641912	SRR12092017	ncRNA-seq	Single			86	1	5		
PRJNA641912	SRR12092018	ncRNA-seq	Single			95	1	4		
PRJNA641912	SRR12092019	ncRNA-seq	Single			76	1	5		
PRJNA662923	SRR12623869	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623870	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623871	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623872	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623873	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623874	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623875	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623876	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623877	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623878	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623879	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623880	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623881	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623882	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623883	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623884	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623885	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623886	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623887	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623888	snRNA-seq	Paired			73	1	6		
PRJNA662923	SRR12623889	snRNA-seq	Paired			71	2	6		
PRJNA662923	SRR12623890	snRNA-seq	Paired			71	2	6		
PRJNA662923	SRR12623891	snRNA-seq	Paired			71	2	6		

PRJNA662923	SRR12623892	snRNA-seq	Paired			71	2	6		
PRJNA662923	SRR12623893	snRNA-seq	Paired			72	1	5		
PRJNA662923	SRR12623894	snRNA-seq	Paired			72	1	5		
PRJNA662923	SRR12623895	snRNA-seq	Paired			72	1	5		
PRJNA662923	SRR12623896	snRNA-seq	Paired			72	1	5		
PRJNA662923	SRR12623897	snRNA-seq	Paired			69	1	5		
PRJNA662923	SRR12623898	snRNA-seq	Paired			69	1	5		
PRJNA662923	SRR12623899	snRNA-seq	Paired			69	1	5		
PRJNA662923	SRR12623900	snRNA-seq	Paired			69	1	5		
PRJNA662923	SRR12623901	snRNA-seq	Paired			70	2	4		
PRJNA662923	SRR12623902	snRNA-seq	Paired			70	2	4		
PRJNA662923	SRR12623903	snRNA-seq	Paired			70	2	4		
PRJNA662923	SRR12623904	snRNA-seq	Paired			70	2	4		
PRJNA662923	SRR12623905	snRNA-seq	Paired			73	1	5		
PRJNA662923	SRR12623906	snRNA-seq	Paired			73	1	5		
PRJNA662923	SRR12623907	snRNA-seq	Paired			73	1	5		
PRJNA662923	SRR12623908	snRNA-seq	Paired			73	1	5		
PRJNA662923	SRR12623909	snRNA-seq	Paired			72	1	6		
PRJNA662923	SRR12623910	snRNA-seq	Paired			72	1	6		
PRJNA662923	SRR12623911	snRNA-seq	Paired			72	1	6		
PRJNA662923	SRR12623912	snRNA-seq	Paired			72	1	6		
PRJNA662923	SRR12623913	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623914	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623915	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623916	snRNA-seq	Paired			74	1	4		
PRJNA662923	SRR12623917	snRNA-seq	Paired			69	1	4		
PRJNA662923	SRR12623918	snRNA-seq	Paired			69	1	4		
PRJNA662923	SRR12623919	snRNA-seq	Paired			69	1	4		
PRJNA662923	SRR12623920	snRNA-seq	Paired			69	1	4		
PRJNA662923	SRR12623921	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623922	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623923	snRNA-seq	Paired			70	2	5		
PRJNA662923	SRR12623924	snRNA-seq	Paired			70	2	5		
PRJNA232671	SRR1063412	RNA-seq	Single	Ctrl	DLPFC/BA9	85		0		
PRJNA232671	SRR1063413	RNA-seq	Single			80		1		
PRJNA232671	SRR1063414	RNA-seq	Single			90		0		
PRJNA210613	SRR1103937	snRNA-seq	Single			80	1	2		
PRJNA210613	SRR1103938	snRNA-seq	Single			78	1	2		
PRJNA210613	SRR1103939	snRNA-seq	Single			52	2	2		

PRJNA210613	SRR1103940	snRNA-seq	Single			74	2	2		
PRJNA210613	SRR1103941	snRNA-seq	Single			83	1	2		
PRJNA210613	SRR1103942	snRNA-seq	Single			78	2	2		
PRJNA434441	SRR6742935	RNA-seq	Single			72	2	2		
PRJNA434441	SRR6742936	RNA-seq	Single			72	2	2		
PRJNA434441	SRR6742939	RNA-seq	Single			65	1	1		
PRJNA434441	SRR6742940	RNA-seq	Single			65	1	1		
PRJNA434441	SRR6742945	RNA-seq	Single			85	1	2		
PRJNA434441	SRR6742946	RNA-seq	Single			85	1	2		
PRJNA434441	SRR6742947	RNA-seq	Single			91	1	2		
PRJNA434441	SRR6742948	RNA-seq	Single			91	1	2		
PRJNA434441	SRR6742949	RNA-seq	Single			96	1	2		
PRJNA434441	SRR6742950	RNA-seq	Single			96	1	2		
PRJNA434441	SRR6742953	RNA-seq	Single			96	2	2		
PRJNA434441	SRR6742954	RNA-seq	Single			96	2	2		
PRJNA434441	SRR6742957	RNA-seq	Single			87	2	2		
PRJNA434441	SRR6742958	RNA-seq	Single			87	2	2		
PRJNA434441	SRR6742963	RNA-seq	Single			89	2	2		
PRJNA434441	SRR6742964	RNA-seq	Single			89	2	2		
PRJNA434441	SRR6742971	RNA-seq	Single			90	1	2		
PRJNA434441	SRR6742972	RNA-seq	Single			90	1	2		
PRJNA434441	SRR6742977	RNA-seq	Single			80	1	2		
PRJNA434441	SRR6742978	RNA-seq	Single			80	1	2		
PRJNA641912	SRR12092020	ncRNA-seq	Single			85	1	2		
PRJNA641912	SRR12092021	ncRNA-seq	Single			74	2	1		
PRJNA641912	SRR12092022	ncRNA-seq	Single			79	2	2		
PRJNA641912	SRR12092023	ncRNA-seq	Single			86	1	2		
PRJNA662923	SRR12623925	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623926	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623927	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623928	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623929	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623930	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623931	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623932	snRNA-seq	Paired			85	1	1		
PRJNA662923	SRR12623933	snRNA-seq	Paired			80	2	1		
PRJNA662923	SRR12623934	snRNA-seq	Paired			80	2	1		
PRJNA662923	SRR12623935	snRNA-seq	Paired			80	2	1		
PRJNA662923	SRR12623936	snRNA-seq	Paired			80	2	1		

PRJNA662923	SRR12623937	snRNA-seq	Paired			81	1	1		
PRJNA662923	SRR12623938	snRNA-seq	Paired			81	1	1		
PRJNA662923	SRR12623939	snRNA-seq	Paired			81	1	1		
PRJNA662923	SRR12623940	snRNA-seq	Paired			81	1	1		
PRJNA662923	SRR12623941	snRNA-seq	Paired			79	2	1		
PRJNA662923	SRR12623942	snRNA-seq	Paired			79	2	1		
PRJNA662923	SRR12623943	snRNA-seq	Paired			79	2	1		
PRJNA662923	SRR12623944	snRNA-seq	Paired			79	2	1		
PRJNA662923	SRR12623945	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623946	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623947	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623948	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623949	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623950	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623951	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623952	snRNA-seq	Paired			76	1	1		
PRJNA662923	SRR12623953	snRNA-seq	Paired			82	2	1		
PRJNA662923	SRR12623954	snRNA-seq	Paired			82	2	1		
PRJNA662923	SRR12623955	snRNA-seq	Paired			82	2	1		
PRJNA662923	SRR12623956	snRNA-seq	Paired			82	2	1		
PRJNA662923	SRR12623957	snRNA-seq	Paired			84	1	1		
PRJNA662923	SRR12623958	snRNA-seq	Paired			84	1	1		
PRJNA662923	SRR12623959	snRNA-seq	Paired			84	1	1		
PRJNA662923	SRR12623960	snRNA-seq	Paired			84	1	1		
PRJNA283498	SRR2015757	RNA-seq	Paired	PD	DLPFC/BA9	74	1	1	7.2	1
PRJNA283498	SRR2015758	RNA-seq	Paired			83	1	3	7.1	1
PRJNA283498	SRR2015759	RNA-seq	Paired			80	1		8.5	0
PRJNA283498	SRR2015760	RNA-seq	Paired			83	1	3	7.3	1
PRJNA283498	SRR2015761	RNA-seq	Paired			80	1	2	8.4	0
PRJNA283498	SRR2015762	RNA-seq	Paired			84	1	2	6.7	1
PRJNA283498	SRR2015763	RNA-seq	Paired			88	1	2	8.1	1
PRJNA283498	SRR2015764	RNA-seq	Paired			81	1	1	5.8	0
PRJNA283498	SRR2015765	RNA-seq	Paired			77	1	3	6.1	0
PRJNA283498	SRR2015766	RNA-seq	Paired			64	1	2	8.1	0
PRJNA283498	SRR2015767	RNA-seq	Paired			85	1	3	6.1	1
PRJNA283498	SRR2015768	RNA-seq	Paired			94	1	3	6.4	
PRJNA283498	SRR2015769	RNA-seq	Paired			80	1	1	6.5	1
PRJNA283498	SRR2015770	RNA-seq	Paired			85	1	1	6.9	0
PRJNA283498	SRR2015771	RNA-seq	Paired			75	1	1	8	
PRJNA283498	SRR2015772	RNA-seq	Paired			74	1	2	7.1	

PRJNA283498	SRR2015773	RNA-seq	Paired			89	1	2	6.1	0
PRJNA283498	SRR2015774	RNA-seq	Paired			66	1	1	7.7	0
PRJNA283498	SRR2015775	RNA-seq	Paired			65	1	1	6.8	
PRJNA283498	SRR2015776	RNA-seq	Paired			85	1		6.1	
PRJNA283498	SRR2015777	RNA-seq	Paired			64	1	2	7.8	1
PRJNA283498	SRR2015778	RNA-seq	Paired			64	1	1	6.1	0
PRJNA283498	SRR2015779	RNA-seq	Paired			75	1	2	6.9	1
PRJNA283498	SRR2015780	RNA-seq	Paired			68	1	1	7.3	1
PRJNA283498	SRR2015781	RNA-seq	Paired			95	1	1	7	1
PRJNA283498	SRR2015782	RNA-seq	Paired			74	1	2	7.5	0
PRJNA283498	SRR2015783	RNA-seq	Paired			68	1	2	7.9	
PRJNA283498	SRR2015784	RNA-seq	Paired			79	1	2	6.9	0
PRJNA283498	SRR2015785	RNA-seq	Paired			70	1	3	6.7	0
PRJNA290524	SRR2121909	RNA-seq	Paired			75			6.1	
PRJNA557205	SRR9854771	RNA-seq	Single			77	1		7.6	
PRJNA557205	SRR9854772	RNA-seq	Single			77	1		7.6	
PRJNA557205	SRR9854773	RNA-seq	Single			77	1		7.6	
PRJNA557205	SRR9854774	RNA-seq	Single			77	1		7.6	
PRJNA557205	SRR9854775	RNA-seq	Single			80	1		6.6	
PRJNA557205	SRR9854776	RNA-seq	Single			80	1		6.6	
PRJNA557205	SRR9854777	RNA-seq	Single			80	1		6.6	
PRJNA557205	SRR9854778	RNA-seq	Single			80	1		6.6	
PRJNA557205	SRR9854779	RNA-seq	Single			78	2		6.7	
PRJNA557205	SRR9854780	RNA-seq	Single			78	2		6.7	
PRJNA557205	SRR9854781	RNA-seq	Single			78	2		6.7	
PRJNA557205	SRR9854782	RNA-seq	Single			78	2		6.7	
PRJNA557205	SRR9854783	RNA-seq	Single			73	1		6	
PRJNA557205	SRR9854784	RNA-seq	Single			73	1		6	
PRJNA557205	SRR9854785	RNA-seq	Single			73	1		6	
PRJNA557205	SRR9854786	RNA-seq	Single			73	1		6	
PRJNA557205	SRR9854787	RNA-seq	Single			75	1		7.7	
PRJNA557205	SRR9854788	RNA-seq	Single			75	1		7.7	
PRJNA557205	SRR9854789	RNA-seq	Single			75	1		7.7	
PRJNA557205	SRR9854790	RNA-seq	Single			75	1		7.7	
PRJNA557205	SRR9854791	RNA-seq	Single			71	1		7.5	
PRJNA557205	SRR9854792	RNA-seq	Single			71	1		7.5	
PRJNA557205	SRR9854793	RNA-seq	Single			71	1		7.5	
PRJNA557205	SRR9854794	RNA-seq	Single			71	1		7.5	
PRJNA557205	SRR9854795	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854796	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854797	RNA-seq	Single			78	1		7.6	

PRJNA557205	SRR9854798	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854799	RNA-seq	Single			78	1		6.4	
PRJNA557205	SRR9854800	RNA-seq	Single			78	1		6.4	
PRJNA557205	SRR9854801	RNA-seq	Single			78	1		6.4	
PRJNA557205	SRR9854802	RNA-seq	Single			78	1		6.4	
PRJNA557205	SRR9854803	RNA-seq	Single			81	1		7.4	
PRJNA557205	SRR9854804	RNA-seq	Single			81	1		7.4	
PRJNA557205	SRR9854805	RNA-seq	Single			81	1		7.4	
PRJNA557205	SRR9854806	RNA-seq	Single			81	1		7.4	
PRJNA557205	SRR9854807	RNA-seq	Single			72	1		7.9	
PRJNA557205	SRR9854808	RNA-seq	Single			72	1		7.9	
PRJNA557205	SRR9854809	RNA-seq	Single			72	1		7.9	
PRJNA557205	SRR9854810	RNA-seq	Single			72	1		7.9	
PRJNA557205	SRR9854811	RNA-seq	Single			82	1		6.7	
PRJNA557205	SRR9854812	RNA-seq	Single			82	1		6.7	
PRJNA557205	SRR9854813	RNA-seq	Single			82	1		6.7	
PRJNA557205	SRR9854814	RNA-seq	Single			82	1		6.7	
PRJNA557205	SRR9854815	RNA-seq	Single			70	1		6.5	
PRJNA557205	SRR9854816	RNA-seq	Single			70	1		6.5	
PRJNA557205	SRR9854817	RNA-seq	Single			70	1		6.5	
PRJNA557205	SRR9854818	RNA-seq	Single			70	1		6.5	
PRJNA557205	SRR9854819	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854820	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854821	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854822	RNA-seq	Single			78	1		7.6	
PRJNA557205	SRR9854823	RNA-seq	Single			67	1		8.1	
PRJNA557205	SRR9854824	RNA-seq	Single			67	1		8.1	
PRJNA557205	SRR9854825	RNA-seq	Single			67	1		8.1	
PRJNA557205	SRR9854826	RNA-seq	Single			67	1		8.1	
PRJNA557205	SRR9854827	RNA-seq	Single			78	1		7	
PRJNA557205	SRR9854828	RNA-seq	Single			78	1		7	
PRJNA557205	SRR9854829	RNA-seq	Single			78	1		7	
PRJNA557205	SRR9854830	RNA-seq	Single			78	1		7	
PRJNA557205	SRR9854831	RNA-seq	Single			69	1		6.2	
PRJNA557205	SRR9854832	RNA-seq	Single			69	1		6.2	
PRJNA557205	SRR9854833	RNA-seq	Single			69	1		6.2	
PRJNA557205	SRR9854834	RNA-seq	Single			69	1		6.2	
PRJNA557205	SRR9854835	RNA-seq	Single			83	2		6.9	
PRJNA557205	SRR9854836	RNA-seq	Single			83	2		6.9	
PRJNA557205	SRR9854837	RNA-seq	Single			83	2		6.9	
PRJNA557205	SRR9854838	RNA-seq	Single			83	2		6.9	

PRJNA557205	SRR9854839	RNA-seq	Single			81	1		7.2	
PRJNA557205	SRR9854840	RNA-seq	Single			81	1		7.2	
PRJNA557205	SRR9854841	RNA-seq	Single			81	1		7.2	
PRJNA557205	SRR9854842	RNA-seq	Single			81	1		7.2	
PRJNA557205	SRR9854843	RNA-seq	Single			74	1		6.9	
PRJNA557205	SRR9854844	RNA-seq	Single			74	1		6.9	
PRJNA557205	SRR9854845	RNA-seq	Single			74	1		6.9	
PRJNA557205	SRR9854846	RNA-seq	Single			74	1		6.9	
PRJNA557205	SRR9854847	RNA-seq	Single			66	2		8.2	
PRJNA557205	SRR9854848	RNA-seq	Single			66	2		8.2	
PRJNA557205	SRR9854849	RNA-seq	Single			66	2		8.2	
PRJNA557205	SRR9854850	RNA-seq	Single			66	2		8.2	
PRJNA557205	SRR9854851	RNA-seq	Single			73	1		6.7	
PRJNA557205	SRR9854852	RNA-seq	Single			73	1		6.7	
PRJNA557205	SRR9854853	RNA-seq	Single			73	1		6.7	
PRJNA557205	SRR9854854	RNA-seq	Single			73	1		6.7	
PRJNA557205	SRR9854855	RNA-seq	Single			81	1		7.9	
PRJNA557205	SRR9854856	RNA-seq	Single			81	1		7.9	
PRJNA557205	SRR9854857	RNA-seq	Single			81	1		7.9	
PRJNA557205	SRR9854858	RNA-seq	Single			81	1		7.9	
PRJNA557205	SRR9854859	RNA-seq	Single			72	1		7.7	
PRJNA557205	SRR9854860	RNA-seq	Single			72	1		7.7	
PRJNA557205	SRR9854861	RNA-seq	Single			72	1		7.7	
PRJNA557205	SRR9854862	RNA-seq	Single			72	1		7.7	
PRJNA557205	SRR9854863	RNA-seq	Single			77	1		6.8	
PRJNA557205	SRR9854864	RNA-seq	Single			77	1		6.8	
PRJNA557205	SRR9854865	RNA-seq	Single			77	1		6.8	
PRJNA557205	SRR9854866	RNA-seq	Single			77	1		6.8	
PRJNA283498	SRR2015713	RNA-seq	Paired	Ctrl	DLPFC/BA9	73	1	3	7.7	
PRJNA283498	SRR2015714	RNA-seq	Paired			91	1	2	7.9	
PRJNA283498	SRR2015715	RNA-seq	Paired			82	1	3	8.6	
PRJNA283498	SRR2015716	RNA-seq	Paired			97	1	3	9.1	
PRJNA283498	SRR2015717	RNA-seq	Paired			86	1	1	8.6	
PRJNA283498	SRR2015718	RNA-seq	Paired			91	1	4	8.7	
PRJNA283498	SRR2015719	RNA-seq	Paired			81	1	1	6	
PRJNA283498	SRR2015720	RNA-seq	Paired			79	1	1	8.4	
PRJNA283498	SRR2015721	RNA-seq	Paired			63	1	2	6.5	
PRJNA283498	SRR2015722	RNA-seq	Paired			66	1		7.1	
PRJNA283498	SRR2015723	RNA-seq	Paired			69	1		7.8	
PRJNA283498	SRR2015724	RNA-seq	Paired			79	1		8	
PRJNA283498	SRR2015725	RNA-seq	Paired			61	1	1	8.2	

PRJNA283498	SRR2015726	RNA-seq	Paired			58	1	0	8.4	
PRJNA283498	SRR2015727	RNA-seq	Paired			70	1	1	8.2	
PRJNA283498	SRR2015728	RNA-seq	Paired			66	1	1	8.5	
PRJNA283498	SRR2015729	RNA-seq	Paired			73	1	1	7.5	
PRJNA283498	SRR2015730	RNA-seq	Paired			60	1		7.9	
PRJNA283498	SRR2015731	RNA-seq	Paired			76	1	1	7.3	
PRJNA283498	SRR2015732	RNA-seq	Paired			61	1	0	7.8	
PRJNA283498	SRR2015733	RNA-seq	Paired			62	1	1	6.6	
PRJNA283498	SRR2015734	RNA-seq	Paired			69	1	1	8.7	
PRJNA283498	SRR2015735	RNA-seq	Paired			61	1	1	8.1	
PRJNA283498	SRR2015736	RNA-seq	Paired			88	1	2	7.1	
PRJNA283498	SRR2015737	RNA-seq	Paired			93	1		6.4	
PRJNA283498	SRR2015738	RNA-seq	Paired			53	1	0	7.3	
PRJNA283498	SRR2015739	RNA-seq	Paired			57	1	1	8.3	
PRJNA283498	SRR2015740	RNA-seq	Paired			46	1	0	7.6	
PRJNA283498	SRR2015741	RNA-seq	Paired			57	1	1	7.7	
PRJNA283498	SRR2015742	RNA-seq	Paired			80	1	1	7.3	
PRJNA283498	SRR2015743	RNA-seq	Paired			74	1	2	8.5	
PRJNA283498	SRR2015744	RNA-seq	Paired			69	1	1	8.4	
PRJNA283498	SRR2015745	RNA-seq	Paired			76	1	1	7.5	
PRJNA283498	SRR2015746	RNA-seq	Paired			78	1	2	7.6	
PRJNA283498	SRR2015747	RNA-seq	Paired			87	1	4	8.7	
PRJNA283498	SRR2015748	RNA-seq	Paired			86	1	2	8.7	
PRJNA283498	SRR2015749	RNA-seq	Paired			54	1		8.3	
PRJNA283498	SRR2015750	RNA-seq	Paired			68	1	0	6.3	
PRJNA283498	SRR2015751	RNA-seq	Paired			52	1	0	7.4	
PRJNA283498	SRR2015752	RNA-seq	Paired			46	1	0	8.2	
PRJNA283498	SRR2015753	RNA-seq	Paired			55	1		7.6	
PRJNA283498	SRR2015754	RNA-seq	Paired			57	1	1	7.8	
PRJNA283498	SRR2015755	RNA-seq	Paired			66	1	1	8.4	
PRJNA283498	SRR2015756	RNA-seq	Paired			64	1	1	8.7	
PRJNA290524	SRR2121907	RNA-seq	Paired			63			6.5	
PRJNA557205	SRR9854867	RNA-seq	Single			79	2		7.5	
PRJNA557205	SRR9854868	RNA-seq	Single			79	2		7.5	
PRJNA557205	SRR9854869	RNA-seq	Single			79	2		7.5	
PRJNA557205	SRR9854870	RNA-seq	Single			79	2		7.5	
PRJNA557205	SRR9854871	RNA-seq	Single			82	1		8.3	
PRJNA557205	SRR9854872	RNA-seq	Single			82	1		8.3	
PRJNA557205	SRR9854873	RNA-seq	Single			82	1		8.3	
PRJNA557205	SRR9854874	RNA-seq	Single			82	1		8.3	
PRJNA557205	SRR9854875	RNA-seq	Single			70	1		7.7	

PRJNA557205	SRR9854876	RNA-seq	Single			70	1		7.7	
PRJNA557205	SRR9854877	RNA-seq	Single			70	1		7.7	
PRJNA557205	SRR9854878	RNA-seq	Single			70	1		7.7	
PRJNA557205	SRR9854879	RNA-seq	Single			77	1		7.1	
PRJNA557205	SRR9854880	RNA-seq	Single			77	1		7.1	
PRJNA557205	SRR9854881	RNA-seq	Single			77	1		7.1	
PRJNA557205	SRR9854882	RNA-seq	Single			77	1		7.1	
PRJNA557205	SRR9854883	RNA-seq	Single			73	1		7.8	
PRJNA557205	SRR9854884	RNA-seq	Single			73	1		7.8	
PRJNA557205	SRR9854885	RNA-seq	Single			73	1		7.8	
PRJNA557205	SRR9854886	RNA-seq	Single			73	1		7.8	
PRJNA557205	SRR9854887	RNA-seq	Single			73	1		7.1	
PRJNA557205	SRR9854888	RNA-seq	Single			73	1		7.1	
PRJNA557205	SRR9854889	RNA-seq	Single			73	1		7.1	
PRJNA557205	SRR9854890	RNA-seq	Single			73	1		7.1	
PRJNA557205	SRR9854891	RNA-seq	Single			68	1		8.1	
PRJNA557205	SRR9854892	RNA-seq	Single			68	1		8.1	
PRJNA557205	SRR9854893	RNA-seq	Single			68	1		8.1	
PRJNA557205	SRR9854894	RNA-seq	Single			68	1		8.1	
PRJNA557205	SRR9854895	RNA-seq	Single			68	1		7.6	
PRJNA557205	SRR9854896	RNA-seq	Single			68	1		7.6	
PRJNA557205	SRR9854897	RNA-seq	Single			68	1		7.6	
PRJNA557205	SRR9854898	RNA-seq	Single			68	1		7.6	
PRJNA557205	SRR9854899	RNA-seq	Single			84	1		8.2	
PRJNA557205	SRR9854900	RNA-seq	Single			84	1		8.2	
PRJNA557205	SRR9854901	RNA-seq	Single			84	1		8.2	
PRJNA557205	SRR9854902	RNA-seq	Single			84	1		8.2	
PRJNA557205	SRR9854903	RNA-seq	Single			71	1		8.8	
PRJNA557205	SRR9854904	RNA-seq	Single			71	1		8.8	
PRJNA557205	SRR9854905	RNA-seq	Single			71	1		8.8	
PRJNA557205	SRR9854906	RNA-seq	Single			71	1		8.8	
PRJNA557205	SRR9854907	RNA-seq	Single			83	1		7.3	
PRJNA557205	SRR9854908	RNA-seq	Single			83	1		7.3	
PRJNA557205	SRR9854909	RNA-seq	Single			83	1		7.3	
PRJNA557205	SRR9854910	RNA-seq	Single			83	1		7.3	
PRJNA557205	SRR9854911	RNA-seq	Single			70	1		8	
PRJNA557205	SRR9854912	RNA-seq	Single			70	1		8	
PRJNA557205	SRR9854913	RNA-seq	Single			70	1		8	
PRJNA557205	SRR9854914	RNA-seq	Single			70	1		8	
PRJNA678865	SRR13071573	RNA-seq	Paired	AD	PBMC	70	2			
PRJNA678865	SRR13071574	RNA-seq	Paired			73	2			

PRJNA678865	SRR13071575	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071576	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071577	RNA-seq	Paired			75	2			
PRJNA678865	SRR13071585	RNA-seq	Paired			70	2			
PRJNA678865	SRR13071586	RNA-seq	Paired			73	2			
PRJNA678865	SRR13071587	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071588	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071589	RNA-seq	Paired			75	2			
PRJNA667860	SRR12786082	AMPLICON	Paired			66	2			
PRJNA667860	SRR12786083	AMPLICON	Paired			76	1			
PRJNA667860	SRR12786084	AMPLICON	Paired			67	2			
PRJNA667860	SRR12786085	AMPLICON	Paired			67	2			
PRJNA667860	SRR12786086	AMPLICON	Paired			69	2			
PRJNA667860	SRR12786087	AMPLICON	Paired			68	2			
PRJNA667860	SRR12786088	AMPLICON	Paired			68	2			
PRJNA667860	SRR12786089	AMPLICON	Paired			81	2			
PRJNA667860	SRR12786094	AMPLICON	Paired			63	1			
PRJNA667860	SRR12786095	AMPLICON	Paired			78	2			
PRJNA667860	SRR12786096	AMPLICON	Paired			79	2			
PRJNA667860	SRR12786097	AMPLICON	Paired			61	1			
PRJNA678865	SRR13071578	RNA-seq	Paired	Ctrl	PBMC	76	2			
PRJNA678865	SRR13071579	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071584	RNA-seq	Paired			74	1			
PRJNA678865	SRR13071590	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071591	RNA-seq	Paired			76	2			
PRJNA678865	SRR13071592	RNA-seq	Paired			74	1			
PRJNA678865	SRR13071593	RNA-seq	Paired			61	1			
PRJNA678865	SRR13071594	RNA-seq	Paired			58	1			
PRJNA678865	SRR13071595	RNA-seq	Paired			61	1			
PRJNA678865	SRR13071596	RNA-seq	Paired			58	1			
PRJNA667860	SRR12786090	AMPLICON	Paired			72	1			
PRJNA667860	SRR12786091	AMPLICON	Paired			78	1			
PRJNA667860	SRR12786092	AMPLICON	Paired			78	1			
PRJNA667860	SRR12786093	AMPLICON	Paired			76	1			
PRJNA675864	SRR13019398	RNA-seq	Paired	PD	PBMC	62	1			
PRJNA675864	SRR13019399	RNA-seq	Paired			68	1			
PRJNA675864	SRR13019400	RNA-seq	Paired			56	1			
PRJNA675864	SRR13019401	RNA-seq	Paired			67	1			
PRJNA675864	SRR13019402	RNA-seq	Paired			64	2			

PRJNA675864	SRR13019403	RNA-seq	Paired			58	2			
PRJNA675864	SRR13019394	RNA-seq	Paired	Ctrl	PBMC	53	2			
PRJNA675864	SRR13019395	RNA-seq	Paired			61	1			
PRJNA675864	SRR13019396	RNA-seq	Paired			45	1			
PRJNA675864	SRR13019397	RNA-seq	Paired			56	1			
PRJNA675864	SRR13019410	RNA-seq	Paired			60	2			
PRJNA675864	SRR13019411	RNA-seq	Paired			57	1			
PRJNA675864	SRR13019412	RNA-seq	Paired			48	2			
PRJNA675864	SRR13019413	RNA-seq	Paired			59	1			
PRJNA675864	SRR13019414	RNA-seq	Paired			51	1			
PRJNA675864	SRR13019415	RNA-seq	Paired			44	1			
PRJNA675864	SRR13019416	RNA-seq	Paired			50	1			
PRJNA201039	SRR837437	RNA-seq	Single	AD	WB	77	2			
PRJNA201039	SRR837438	RNA-seq	Single			74	1			
PRJNA201039	SRR837439	RNA-seq	Single			68	1			
PRJNA201039	SRR837440	RNA-seq	Single			75	2			
PRJNA201039	SRR837441	RNA-seq	Single			74	1			
PRJNA201039	SRR837442	RNA-seq	Single			76	1			
PRJNA201039	SRR837443	RNA-seq	Single			79	1			
PRJNA201039	SRR837444	RNA-seq	Single			75	2			
PRJNA201039	SRR837445	RNA-seq	Single			77	1			
PRJNA201039	SRR837446	RNA-seq	Single			72	2			
PRJNA201039	SRR837447	RNA-seq	Single			76	1			
PRJNA201039	SRR837448	RNA-seq	Single			61	1			
PRJNA201039	SRR837449	RNA-seq	Single			51	2			
PRJNA201039	SRR837450	RNA-seq	Single			55	2			
PRJNA201039	SRR837451	RNA-seq	Single			70	2			
PRJNA201039	SRR837452	RNA-seq	Single			78	2			
PRJNA201039	SRR837453	RNA-seq	Single			81	1			
PRJNA201039	SRR837454	RNA-seq	Single			72	2			
PRJNA201039	SRR837455	RNA-seq	Single			57	1			
PRJNA201039	SRR837456	RNA-seq	Single			75	1			
PRJNA201039	SRR837457	RNA-seq	Single			70	2			
PRJNA201039	SRR837458	RNA-seq	Single			75	2			
PRJNA201039	SRR837459	RNA-seq	Single			69	2			
PRJNA201039	SRR837460	RNA-seq	Single			68	2			
PRJNA201039	SRR837461	RNA-seq	Single			71	2			
PRJNA201039	SRR837462	RNA-seq	Single			74	1			
PRJNA201039	SRR837463	RNA-seq	Single			72	1			
PRJNA201039	SRR837464	RNA-seq	Single			69	2			

PRJNA201039	SRR837465	RNA-seq	Single			63	2			
PRJNA201039	SRR837466	RNA-seq	Single			80	1			
PRJNA201039	SRR837467	RNA-seq	Single			62	1			
PRJNA201039	SRR837468	RNA-seq	Single			75	1			
PRJNA201039	SRR837469	RNA-seq	Single			56	2			
PRJNA201039	SRR837470	RNA-seq	Single			57	2			
PRJNA201039	SRR837493	RNA-seq	Single			80	1			
PRJNA201039	SRR837494	RNA-seq	Single			81	1			
PRJNA201039	SRR837495	RNA-seq	Single			76	2			
PRJNA201039	SRR837496	RNA-seq	Single			60	2			
PRJNA201039	SRR837497	RNA-seq	Single			63	2			
PRJNA201039	SRR837498	RNA-seq	Single			64	2			
PRJNA201039	SRR837499	RNA-seq	Single			65	1			
PRJNA201039	SRR837500	RNA-seq	Single			74	2			
PRJNA201039	SRR837501	RNA-seq	Single			59	2			
PRJNA201039	SRR837502	RNA-seq	Single			72	2			
PRJNA201039	SRR837503	RNA-seq	Single			63	1			
PRJNA201039	SRR837504	RNA-seq	Single			78	1			
PRJNA201039	SRR837505	RNA-seq	Single			78	1			
PRJNA201039	SRR837506	RNA-seq	Single			79	1			
PRJNA201039	SRR837471	RNA-seq	Single	Ctrl	WB	76	1			
PRJNA201039	SRR837472	RNA-seq	Single			60	2			
PRJNA201039	SRR837473	RNA-seq	Single			60	2			
PRJNA201039	SRR837474	RNA-seq	Single			76	2			
PRJNA201039	SRR837475	RNA-seq	Single			60	2			
PRJNA201039	SRR837476	RNA-seq	Single			83	1			
PRJNA201039	SRR837477	RNA-seq	Single			61	1			
PRJNA201039	SRR837478	RNA-seq	Single			77	1			
PRJNA201039	SRR837479	RNA-seq	Single			63	1			
PRJNA201039	SRR837480	RNA-seq	Single			75	1			
PRJNA201039	SRR837481	RNA-seq	Single			63	1			
PRJNA201039	SRR837482	RNA-seq	Single			64	1			
PRJNA201039	SRR837483	RNA-seq	Single			71	2			
PRJNA201039	SRR837484	RNA-seq	Single			60	2			
PRJNA201039	SRR837485	RNA-seq	Single			75	1			
PRJNA201039	SRR837486	RNA-seq	Single			60	2			
PRJNA201039	SRR837487	RNA-seq	Single			66	2			
PRJNA201039	SRR837488	RNA-seq	Single			60	2			
PRJNA201039	SRR837489	RNA-seq	Single			61	1			
PRJNA201039	SRR837490	RNA-seq	Single			74	2			
PRJNA201039	SRR837491	RNA-seq	Single			71	2			

PRJNA201039	SRR837492	RNA-seq	Single			60	1			
PRJEB8960	ERR833219	RNA-seq	Single	PD	WB	59	2			
PRJEB8960	ERR833220	RNA-seq	Single			59	2			
PRJEB8960	ERR833221	RNA-seq	Single			59	2			
PRJEB8960	ERR833222	RNA-seq	Single			59	2			
PRJEB8960	ERR833227	RNA-seq	Single			53	1			
PRJEB8960	ERR833228	RNA-seq	Single			53	1			
PRJEB8960	ERR833229	RNA-seq	Single			53	1			
PRJEB8960	ERR833230	RNA-seq	Single			53	1			
PRJEB8960	ERR833239	RNA-seq	Single			50	1			
PRJEB8960	ERR833240	RNA-seq	Single			50	1			
PRJEB8960	ERR833241	RNA-seq	Single			50	1			
PRJEB8960	ERR833242	RNA-seq	Single			50	1			
PRJEB8960	ERR833251	RNA-seq	Single			47	1			
PRJEB8960	ERR833252	RNA-seq	Single			47	1			
PRJEB8960	ERR833253	RNA-seq	Single			47	1			
PRJEB8960	ERR833254	RNA-seq	Single			47	1			
PRJEB8960	ERR833255	RNA-seq	Single			50	1			
PRJEB8960	ERR833256	RNA-seq	Single			50	1			
PRJEB8960	ERR833257	RNA-seq	Single			50	1			
PRJEB8960	ERR833258	RNA-seq	Single			50	1			
PRJEB8960	ERR833223	RNA-seq	Single	Ctrl	WB	54	2			
PRJEB8960	ERR833224	RNA-seq	Single			54	2			
PRJEB8960	ERR833225	RNA-seq	Single			54	2			
PRJEB8960	ERR833226	RNA-seq	Single			54	2			
PRJEB8960	ERR833231	RNA-seq	Single			45	1			
PRJEB8960	ERR833232	RNA-seq	Single			45	1			
PRJEB8960	ERR833233	RNA-seq	Single			45	1			
PRJEB8960	ERR833234	RNA-seq	Single			45	1			
PRJEB8960	ERR833235	RNA-seq	Single			57	1			
PRJEB8960	ERR833236	RNA-seq	Single			57	1			
PRJEB8960	ERR833237	RNA-seq	Single			57	1			
PRJEB8960	ERR833238	RNA-seq	Single			57	1			
PRJEB8960	ERR833243	RNA-seq	Single			42	2			
PRJEB8960	ERR833244	RNA-seq	Single			42	2			
PRJEB8960	ERR833245	RNA-seq	Single			42	2			
PRJEB8960	ERR833246	RNA-seq	Single			42	2			
PRJEB8960	ERR833247	RNA-seq	Single			42	1			
PRJEB8960	ERR833248	RNA-seq	Single			42	1			
PRJEB8960	ERR833249	RNA-seq	Single			42	1			
PRJEB8960	ERR833250	RNA-seq	Single			42	1			

Appendix 3 Table provides information of RNA-seq samples with ~150bp read lengths downloaded from SRA.

The SRA publicly accessioned objects available in the meta-data: Project (study): describes the goal of the research effort. Strategy: describes the sequencing strategy. SID (SRR#): a manifest of data file(s) that are derived from sequencing a library.

Type of tissue and the related human disease can be found in the Tissue and Condition column, respectively.

Related physiological traits can be found in Age, Sex, Braak, RIN, and Dementia column.

Project	Strategy	SID	Condition	Tissue	Age	Sex	Braak	RIN	Dementia
PRJNA232671	RNA-seq	SRR1063418	AD	DLPFC/BA9	100		4		
PRJNA232671	RNA-seq	SRR1063419			88		5		
PRJNA232671	RNA-seq	SRR1063420			85		5		
PRJNA210613	RNA-seq	SRR1103937			80	1	2		
PRJNA210613	RNA-seq	SRR1103938			78	1	2		
PRJNA210613	RNA-seq	SRR1103939			52	2	2		
PRJNA210613	RNA-seq	SRR1103940			74	2	2		
PRJNA210613	RNA-seq	SRR1103941			83	1	2		
PRJNA210613	RNA-seq	SRR1103942			78	2	2		
PRJNA210613	RNA-seq	SRR1103943			79	1	5		
PRJNA210613	RNA-seq	SRR1103944			80	2	6		
PRJNA210613	RNA-seq	SRR1103945			67	1	6		
PRJNA210613	RNA-seq	SRR1103946			67	2	7		
PRJNA210613	RNA-seq	SRR1103947			68	2	8		
PRJNA210613	RNA-seq	SRR1103948			72	2	9		
PRJNA232669	RNA-seq	SRR2422948							
PRJNA232669	RNA-seq	SRR2422949							
PRJNA232669	RNA-seq	SRR2422950							
PRJNA232669	RNA-seq	SRR2422951							
PRJNA232669	RNA-seq	SRR2422952							
PRJNA232669	RNA-seq	SRR2422953							
PRJNA434441	RNA-seq	SRR6742935			72	2	2		
PRJNA434441	RNA-seq	SRR6742936			72	2	2		
PRJNA434441	RNA-seq	SRR6742939			65	1	1		
PRJNA434441	RNA-seq	SRR6742940			65	1	1		
PRJNA434441	RNA-seq	SRR6742945			85	1	2		
PRJNA434441	RNA-seq	SRR6742946			85	1	2		
PRJNA434441	RNA-seq	SRR6742947			91	1	2		
PRJNA434441	RNA-seq	SRR6742948			91	1	2		
PRJNA434441	RNA-seq	SRR6742949			96	1	2		
PRJNA434441	RNA-seq	SRR6742950			96	1	2		

PRJNA434441	RNA-seq	SRR6742953			96	2	2		
PRJNA434441	RNA-seq	SRR6742954			96	2	2		
PRJNA434441	RNA-seq	SRR6742957			87	2	2		
PRJNA434441	RNA-seq	SRR6742958			87	2	2		
PRJNA434441	RNA-seq	SRR6742963			89	2	2		
PRJNA434441	RNA-seq	SRR6742964			89	2	2		
PRJNA434441	RNA-seq	SRR6742971			90	1	2		
PRJNA434441	RNA-seq	SRR6742972			90	1	2		
PRJNA434441	RNA-seq	SRR6742977			80	1	2		
PRJNA434441	RNA-seq	SRR6742978			80	1	2		
PRJNA434441	RNA-seq	SRR6742933			88	1	5		
PRJNA434441	RNA-seq	SRR6742934			88	1	5		
PRJNA434441	RNA-seq	SRR6742937			85	1	6		
PRJNA434441	RNA-seq	SRR6742938			85	1	6		
PRJNA434441	RNA-seq	SRR6742967			73	2	6		
PRJNA434441	RNA-seq	SRR6742968			73	2	6		
PRJNA434441	RNA-seq	SRR6742973			96	2	6		
PRJNA434441	RNA-seq	SRR6742974			96	2	6		
PRJNA434441	RNA-seq	SRR6742975			85	2	6		
PRJNA434441	RNA-seq	SRR6742976			85	2	6		
PRJNA434441	RNA-seq	SRR6742981			89	2	6		
PRJNA434441	RNA-seq	SRR6742982			89	2	6		
PRJNA434441	RNA-seq	SRR6742941			98	1	4		
PRJNA434441	RNA-seq	SRR6742942			98	1	4		
PRJNA434441	RNA-seq	SRR6742943			92	1	3		
PRJNA434441	RNA-seq	SRR6742944			92	1	3		
PRJNA434441	RNA-seq	SRR6742951			93	1	4		
PRJNA434441	RNA-seq	SRR6742952			93	1	4		
PRJNA434441	RNA-seq	SRR6742955			86	1	4		
PRJNA434441	RNA-seq	SRR6742956			86	1	4		
PRJNA434441	RNA-seq	SRR6742959			88	2	3		
PRJNA434441	RNA-seq	SRR6742960			88	2	3		
PRJNA434441	RNA-seq	SRR6742961			78	1	3		
PRJNA434441	RNA-seq	SRR6742962			78	1	3		
PRJNA434441	RNA-seq	SRR6742965			90	2	4		
PRJNA434441	RNA-seq	SRR6742966			90	2	4		
PRJNA434441	RNA-seq	SRR6742969			75	2	3		
PRJNA434441	RNA-seq	SRR6742970			75	2	3		
PRJNA434441	RNA-seq	SRR6742979			86	2	3		
PRJNA434441	RNA-seq	SRR6742980			86	2	3		
PRJNA641912	ncRNA-seq	SRR12092014			84	2	5		

PRJNA641912	ncRNA-seq	SRR12092015			78	1	6		
PRJNA641912	ncRNA-seq	SRR12092016			98	2	5		
PRJNA641912	ncRNA-seq	SRR12092017			86	1	5		
PRJNA641912	ncRNA-seq	SRR12092018			95	1	4		
PRJNA641912	ncRNA-seq	SRR12092019			76	1	5		
PRJNA662923	RNA-seq	SRR12623869			74	1	4		
PRJNA662923	RNA-seq	SRR12623870			74	1	4		
PRJNA662923	RNA-seq	SRR12623871			74	1	4		
PRJNA662923	RNA-seq	SRR12623872			74	1	4		
PRJNA662923	RNA-seq	SRR12623873			74	1	4		
PRJNA662923	RNA-seq	SRR12623874			74	1	4		
PRJNA662923	RNA-seq	SRR12623875			74	1	4		
PRJNA662923	RNA-seq	SRR12623876			74	1	4		
PRJNA662923	RNA-seq	SRR12623877			70	2	5		
PRJNA662923	RNA-seq	SRR12623878			70	2	5		
PRJNA662923	RNA-seq	SRR12623879			70	2	5		
PRJNA662923	RNA-seq	SRR12623880			70	2	5		
PRJNA662923	RNA-seq	SRR12623881			73	1	6		
PRJNA662923	RNA-seq	SRR12623882			73	1	6		
PRJNA662923	RNA-seq	SRR12623883			73	1	6		
PRJNA662923	RNA-seq	SRR12623884			73	1	6		
PRJNA662923	RNA-seq	SRR12623885			73	1	6		
PRJNA662923	RNA-seq	SRR12623886			73	1	6		
PRJNA662923	RNA-seq	SRR12623887			73	1	6		
PRJNA662923	RNA-seq	SRR12623888			73	1	6		
PRJNA662923	RNA-seq	SRR12623889			71	2	6		
PRJNA662923	RNA-seq	SRR12623890			71	2	6		
PRJNA662923	RNA-seq	SRR12623891			71	2	6		
PRJNA662923	RNA-seq	SRR12623892			71	2	6		
PRJNA662923	RNA-seq	SRR12623893			72	1	5		
PRJNA662923	RNA-seq	SRR12623894			72	1	5		
PRJNA662923	RNA-seq	SRR12623895			72	1	5		
PRJNA662923	RNA-seq	SRR12623896			72	1	5		
PRJNA662923	RNA-seq	SRR12623897			69	1	5		
PRJNA662923	RNA-seq	SRR12623898			69	1	5		
PRJNA662923	RNA-seq	SRR12623899			69	1	5		
PRJNA662923	RNA-seq	SRR12623900			69	1	5		
PRJNA662923	RNA-seq	SRR12623901			70	2	4		
PRJNA662923	RNA-seq	SRR12623902			70	2	4		
PRJNA662923	RNA-seq	SRR12623903			70	2	4		
PRJNA662923	RNA-seq	SRR12623904			70	2	4		

PRJNA662923	RNA-seq	SRR12623905			73	1	5		
PRJNA662923	RNA-seq	SRR12623906			73	1	5		
PRJNA662923	RNA-seq	SRR12623907			73	1	5		
PRJNA662923	RNA-seq	SRR12623908			73	1	5		
PRJNA662923	RNA-seq	SRR12623909			72	1	6		
PRJNA662923	RNA-seq	SRR12623910			72	1	6		
PRJNA662923	RNA-seq	SRR12623911			72	1	6		
PRJNA662923	RNA-seq	SRR12623912			72	1	6		
PRJNA662923	RNA-seq	SRR12623913			74	1	4		
PRJNA662923	RNA-seq	SRR12623914			74	1	4		
PRJNA662923	RNA-seq	SRR12623915			74	1	4		
PRJNA662923	RNA-seq	SRR12623916			74	1	4		
PRJNA662923	RNA-seq	SRR12623917			69	1	4		
PRJNA662923	RNA-seq	SRR12623918			69	1	4		
PRJNA662923	RNA-seq	SRR12623919			69	1	4		
PRJNA662923	RNA-seq	SRR12623920			69	1	4		
PRJNA662923	RNA-seq	SRR12623921			70	2	5		
PRJNA662923	RNA-seq	SRR12623922			70	2	5		
PRJNA662923	RNA-seq	SRR12623923			70	2	5		
PRJNA662923	RNA-seq	SRR12623924			70	2	5		
PRJNA232671	RNA-seq	SRR1063412	Ctrl	DLPFC/BA9	85		0		
PRJNA232671	RNA-seq	SRR1063413			80		1		
PRJNA232671	RNA-seq	SRR1063414			90		0		
PRJNA232669	RNA-seq	SRR2422943							
PRJNA232669	RNA-seq	SRR2422944							
PRJNA232669	RNA-seq	SRR2422945							
PRJNA232669	RNA-seq	SRR2422946							
PRJNA232669	RNA-seq	SRR2422947							
PRJNA662923	RNA-seq	SRR12623925			85	1	1		
PRJNA662923	RNA-seq	SRR12623926			85	1	1		
PRJNA662923	RNA-seq	SRR12623927			85	1	1		
PRJNA662923	RNA-seq	SRR12623928			85	1	1		
PRJNA662923	RNA-seq	SRR12623929			85	1	1		
PRJNA662923	RNA-seq	SRR12623930			85	1	1		
PRJNA662923	RNA-seq	SRR12623931			85	1	1		
PRJNA662923	RNA-seq	SRR12623932			85	1	1		
PRJNA662923	RNA-seq	SRR12623933			80	2	1		
PRJNA662923	RNA-seq	SRR12623934			80	2	1		
PRJNA662923	RNA-seq	SRR12623935			80	2	1		
PRJNA662923	RNA-seq	SRR12623936			80	2	1		
PRJNA662923	RNA-seq	SRR12623937			81	1	1		

PRJNA662923	RNA-seq	SRR12623938			81	1	1		
PRJNA662923	RNA-seq	SRR12623939			81	1	1		
PRJNA662923	RNA-seq	SRR12623940			81	1	1		
PRJNA662923	RNA-seq	SRR12623941			79	2	1		
PRJNA662923	RNA-seq	SRR12623942			79	2	1		
PRJNA662923	RNA-seq	SRR12623943			79	2	1		
PRJNA662923	RNA-seq	SRR12623944			79	2	1		
PRJNA662923	RNA-seq	SRR12623945			76	1	1		
PRJNA662923	RNA-seq	SRR12623946			76	1	1		
PRJNA662923	RNA-seq	SRR12623947			76	1	1		
PRJNA662923	RNA-seq	SRR12623948			76	1	1		
PRJNA662923	RNA-seq	SRR12623949			76	1	1		
PRJNA662923	RNA-seq	SRR12623950			76	1	1		
PRJNA662923	RNA-seq	SRR12623951			76	1	1		
PRJNA662923	RNA-seq	SRR12623952			76	1	1		
PRJNA662923	RNA-seq	SRR12623953			82	2	1		
PRJNA662923	RNA-seq	SRR12623954			82	2	1		
PRJNA662923	RNA-seq	SRR12623955			82	2	1		
PRJNA662923	RNA-seq	SRR12623956			82	2	1		
PRJNA662923	RNA-seq	SRR12623957			84	1	1		
PRJNA662923	RNA-seq	SRR12623958			84	1	1		
PRJNA662923	RNA-seq	SRR12623959			84	1	1		
PRJNA662923	RNA-seq	SRR12623960			84	1	1		
PRJNA143627	RNA-seq	SRR306841				1			
PRJNA185476	RNA-seq	SRR646457				2			
PRJNA185476	RNA-seq	SRR646458				2			
PRJNA185476	RNA-seq	SRR646459				2			
PRJNA185476	RNA-seq	SRR646460				2			
PRJEB5212	RNA-seq	ERR409878			43	1			
PRJEB5212	RNA-seq	ERR409879			80	1			
PRJEB5212	RNA-seq	ERR409882			53	1			
PRJEB5212	RNA-seq	ERR409883			46	1			
PRJEB5212	RNA-seq	ERR409887			76	1			
PRJEB5212	RNA-seq	ERR409888			57	1			
PRJEB5212	RNA-seq	ERR409891			79	1			
PRJEB5212	RNA-seq	ERR409893			40	1			
PRJEB5212	RNA-seq	ERR409896			57	1			
PRJEB5212	RNA-seq	ERR409898			93	1			
PRJEB5212	RNA-seq	ERR409900			44	1			
PRJNA283498	RNA-seq	SRR2015757	PD	DLPFC/BA9	74	1	1	7.2	1
PRJNA283498	RNA-seq	SRR2015758			83	1	3	7.1	1

PRJNA283498	RNA-seq	SRR2015759			80	1		8.5	0
PRJNA283498	RNA-seq	SRR2015760			83	1	3	7.3	1
PRJNA283498	RNA-seq	SRR2015761			80	1	2	8.4	0
PRJNA283498	RNA-seq	SRR2015762			84	1	2	6.7	1
PRJNA283498	RNA-seq	SRR2015763			88	1	2	8.1	1
PRJNA283498	RNA-seq	SRR2015764			81	1	1	5.8	0
PRJNA283498	RNA-seq	SRR2015765			77	1	3	6.1	0
PRJNA283498	RNA-seq	SRR2015766			64	1	2	8.1	0
PRJNA283498	RNA-seq	SRR2015767			85	1	3	6.1	1
PRJNA283498	RNA-seq	SRR2015768			94	1	3	6.4	
PRJNA283498	RNA-seq	SRR2015769			80	1	1	6.5	1
PRJNA283498	RNA-seq	SRR2015770			85	1	1	6.9	0
PRJNA283498	RNA-seq	SRR2015771			75	1	1	8	
PRJNA283498	RNA-seq	SRR2015772			74	1	2	7.1	
PRJNA283498	RNA-seq	SRR2015773			89	1	2	6.1	0
PRJNA283498	RNA-seq	SRR2015774			66	1	1	7.7	0
PRJNA283498	RNA-seq	SRR2015775			65	1	1	6.8	
PRJNA283498	RNA-seq	SRR2015776			85	1		6.1	
PRJNA283498	RNA-seq	SRR2015777			64	1	2	7.8	1
PRJNA283498	RNA-seq	SRR2015778			64	1	1	6.1	0
PRJNA283498	RNA-seq	SRR2015779			75	1	2	6.9	1
PRJNA283498	RNA-seq	SRR2015780			68	1	1	7.3	1
PRJNA283498	RNA-seq	SRR2015781			95	1	1	7	1
PRJNA283498	RNA-seq	SRR2015782			74	1	2	7.5	0
PRJNA283498	RNA-seq	SRR2015783			68	1	2	7.9	
PRJNA283498	RNA-seq	SRR2015784			79	1	2	6.9	0
PRJNA283498	RNA-seq	SRR2015785			70	1	3	6.7	0
PRJNA557205	RNA-seq	SRR9854771			77	1		7.6	
PRJNA557205	RNA-seq	SRR9854772			77	1		7.6	
PRJNA557205	RNA-seq	SRR9854773			77	1		7.6	
PRJNA557205	RNA-seq	SRR9854774			77	1		7.6	
PRJNA557205	RNA-seq	SRR9854775			80	1		6.6	
PRJNA557205	RNA-seq	SRR9854776			80	1		6.6	
PRJNA557205	RNA-seq	SRR9854777			80	1		6.6	
PRJNA557205	RNA-seq	SRR9854778			80	1		6.6	
PRJNA557205	RNA-seq	SRR9854779			78	2		6.7	
PRJNA557205	RNA-seq	SRR9854780			78	2		6.7	
PRJNA557205	RNA-seq	SRR9854781			78	2		6.7	
PRJNA557205	RNA-seq	SRR9854782			78	2		6.7	
PRJNA557205	RNA-seq	SRR9854783			73	1		6	
PRJNA557205	RNA-seq	SRR9854784			73	1		6	

PRJNA557205	RNA-seq	SRR9854785			73	1		6	
PRJNA557205	RNA-seq	SRR9854786			73	1		6	
PRJNA557205	RNA-seq	SRR9854787			75	1		7.7	
PRJNA557205	RNA-seq	SRR9854788			75	1		7.7	
PRJNA557205	RNA-seq	SRR9854789			75	1		7.7	
PRJNA557205	RNA-seq	SRR9854790			75	1		7.7	
PRJNA557205	RNA-seq	SRR9854791			71	1		7.5	
PRJNA557205	RNA-seq	SRR9854792			71	1		7.5	
PRJNA557205	RNA-seq	SRR9854793			71	1		7.5	
PRJNA557205	RNA-seq	SRR9854794			71	1		7.5	
PRJNA557205	RNA-seq	SRR9854795			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854796			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854797			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854798			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854799			78	1		6.4	
PRJNA557205	RNA-seq	SRR9854800			78	1		6.4	
PRJNA557205	RNA-seq	SRR9854801			78	1		6.4	
PRJNA557205	RNA-seq	SRR9854802			78	1		6.4	
PRJNA557205	RNA-seq	SRR9854803			81	1		7.4	
PRJNA557205	RNA-seq	SRR9854804			81	1		7.4	
PRJNA557205	RNA-seq	SRR9854805			81	1		7.4	
PRJNA557205	RNA-seq	SRR9854806			81	1		7.4	
PRJNA557205	RNA-seq	SRR9854807			72	1		7.9	
PRJNA557205	RNA-seq	SRR9854808			72	1		7.9	
PRJNA557205	RNA-seq	SRR9854809			72	1		7.9	
PRJNA557205	RNA-seq	SRR9854810			72	1		7.9	
PRJNA557205	RNA-seq	SRR9854811			82	1		6.7	
PRJNA557205	RNA-seq	SRR9854812			82	1		6.7	
PRJNA557205	RNA-seq	SRR9854813			82	1		6.7	
PRJNA557205	RNA-seq	SRR9854814			82	1		6.7	
PRJNA557205	RNA-seq	SRR9854815			70	1		6.5	
PRJNA557205	RNA-seq	SRR9854816			70	1		6.5	
PRJNA557205	RNA-seq	SRR9854817			70	1		6.5	
PRJNA557205	RNA-seq	SRR9854818			70	1		6.5	
PRJNA557205	RNA-seq	SRR9854819			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854820			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854821			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854822			78	1		7.6	
PRJNA557205	RNA-seq	SRR9854823			67	1		8.1	
PRJNA557205	RNA-seq	SRR9854824			67	1		8.1	
PRJNA557205	RNA-seq	SRR9854825			67	1		8.1	

PRJNA557205	RNA-seq	SRR9854826			67	1		8.1	
PRJNA557205	RNA-seq	SRR9854827			78	1		7	
PRJNA557205	RNA-seq	SRR9854828			78	1		7	
PRJNA557205	RNA-seq	SRR9854829			78	1		7	
PRJNA557205	RNA-seq	SRR9854830			78	1		7	
PRJNA557205	RNA-seq	SRR9854831			69	1		6.2	
PRJNA557205	RNA-seq	SRR9854832			69	1		6.2	
PRJNA557205	RNA-seq	SRR9854833			69	1		6.2	
PRJNA557205	RNA-seq	SRR9854834			69	1		6.2	
PRJNA557205	RNA-seq	SRR9854835			83	2		6.9	
PRJNA557205	RNA-seq	SRR9854836			83	2		6.9	
PRJNA557205	RNA-seq	SRR9854837			83	2		6.9	
PRJNA557205	RNA-seq	SRR9854838			83	2		6.9	
PRJNA557205	RNA-seq	SRR9854839			81	1		7.2	
PRJNA557205	RNA-seq	SRR9854840			81	1		7.2	
PRJNA557205	RNA-seq	SRR9854841			81	1		7.2	
PRJNA557205	RNA-seq	SRR9854842			81	1		7.2	
PRJNA557205	RNA-seq	SRR9854843			74	1		6.9	
PRJNA557205	RNA-seq	SRR9854844			74	1		6.9	
PRJNA557205	RNA-seq	SRR9854845			74	1		6.9	
PRJNA557205	RNA-seq	SRR9854846			74	1		6.9	
PRJNA557205	RNA-seq	SRR9854847			66	2		8.2	
PRJNA557205	RNA-seq	SRR9854848			66	2		8.2	
PRJNA557205	RNA-seq	SRR9854849			66	2		8.2	
PRJNA557205	RNA-seq	SRR9854850			66	2		8.2	
PRJNA557205	RNA-seq	SRR9854851			73	1		6.7	
PRJNA557205	RNA-seq	SRR9854852			73	1		6.7	
PRJNA557205	RNA-seq	SRR9854853			73	1		6.7	
PRJNA557205	RNA-seq	SRR9854854			73	1		6.7	
PRJNA557205	RNA-seq	SRR9854855			81	1		7.9	
PRJNA557205	RNA-seq	SRR9854856			81	1		7.9	
PRJNA557205	RNA-seq	SRR9854857			81	1		7.9	
PRJNA557205	RNA-seq	SRR9854858			81	1		7.9	
PRJNA557205	RNA-seq	SRR9854859			72	1		7.7	
PRJNA557205	RNA-seq	SRR9854860			72	1		7.7	
PRJNA557205	RNA-seq	SRR9854861			72	1		7.7	
PRJNA557205	RNA-seq	SRR9854862			72	1		7.7	
PRJNA557205	RNA-seq	SRR9854863			77	1		6.8	
PRJNA557205	RNA-seq	SRR9854864			77	1		6.8	
PRJNA557205	RNA-seq	SRR9854865			77	1		6.8	
PRJNA557205	RNA-seq	SRR9854866			77	1		6.8	

PRJNA295431	miRNA-seq	SRR2353417			74	1		7.7	1
PRJNA295431	miRNA-seq	SRR2353418			83	1		6.9	1
PRJNA295431	miRNA-seq	SRR2353419			80	1		8.5	0
PRJNA295431	miRNA-seq	SRR2353420			83	1		8.1	1
PRJNA295431	miRNA-seq	SRR2353421			80	1		7.7	0
PRJNA295431	miRNA-seq	SRR2353422			84	1		7.2	1
PRJNA295431	miRNA-seq	SRR2353423			88	1		7.5	1
PRJNA295431	miRNA-seq	SRR2353424			81	1		6.4	0
PRJNA295431	miRNA-seq	SRR2353425			77	1		7.2	0
PRJNA295431	miRNA-seq	SRR2353426			64	1		8	0
PRJNA295431	miRNA-seq	SRR2353427			85	1		8.2	1
PRJNA295431	miRNA-seq	SRR2353428			94	1		7.1	0
PRJNA295431	miRNA-seq	SRR2353429			80	1		7.9	1
PRJNA295431	miRNA-seq	SRR2353430			85	1		7.2	0
PRJNA295431	miRNA-seq	SRR2353431			75	1		8	0
PRJNA295431	miRNA-seq	SRR2353432			74	1		6.2	0
PRJNA295431	miRNA-seq	SRR2353433			89	1		6.8	0
PRJNA295431	miRNA-seq	SRR2353434			66	1		8.2	0
PRJNA295431	miRNA-seq	SRR2353435			65	1		8	0
PRJNA295431	miRNA-seq	SRR2353436			85	1		5.5	0
PRJNA295431	miRNA-seq	SRR2353437			64	1		7.8	1
PRJNA295431	miRNA-seq	SRR2353438			64	1		6.1	0
PRJNA295431	miRNA-seq	SRR2353439			75	1		6.9	1
PRJNA295431	miRNA-seq	SRR2353440			68	1		7.3	1
PRJNA295431	miRNA-seq	SRR2353441			95	1		7	1
PRJNA295431	miRNA-seq	SRR2353442			74	1		7.5	0
PRJNA295431	miRNA-seq	SRR2353443			68	1		7.9	0
PRJNA295431	miRNA-seq	SRR2353444			79	1		6.9	0
PRJNA295431	miRNA-seq	SRR2353445			70	1		6.7	0
PRJNA174763	RNA-seq	SRR563541			78	1		9.1	
PRJNA174763	RNA-seq	SRR563542			73	1		8.7	
PRJNA174763	RNA-seq	SRR563543			83	1		7.9	
PRJNA174763	RNA-seq	SRR563544			80	1		8.7	
PRJNA174763	RNA-seq	SRR563545			74	1		8.6	
PRJNA174763	RNA-seq	SRR563546			76	1		6	
PRJNA174763	RNA-seq	SRR563547			82	2		7	
PRJNA174763	RNA-seq	SRR563548			81	1		7.3	
PRJNA174763	RNA-seq	SRR563549			77	2		8.2	
PRJNA174763	RNA-seq	SRR563550			74	1		8.5	
PRJNA174763	RNA-seq	SRR563551			76	1		7.9	
PRJNA174763	RNA-seq	SRR563552			78	2		7.9	

PRJNA174763	RNA-seq	SRR563553			79	1		7.8	
PRJNA174763	RNA-seq	SRR563554			83	2		7.1	
PRJNA174763	RNA-seq	SRR563555			75	1		8.6	
PRJNA174763	RNA-seq	SRR563556			77	1		8.4	
PRJNA174763	RNA-seq	SRR563557			82	2		8.7	
PRJNA283498	RNA-seq	SRR2015713	Ctrl	DLPFC/BA9	73	1	3	7.7	
PRJNA283498	RNA-seq	SRR2015714			91	1	2	7.9	
PRJNA283498	RNA-seq	SRR2015715			82	1	3	8.6	
PRJNA283498	RNA-seq	SRR2015716			97	1	3	9.1	
PRJNA283498	RNA-seq	SRR2015717			86	1	1	8.6	
PRJNA283498	RNA-seq	SRR2015718			91	1	4	8.7	
PRJNA283498	RNA-seq	SRR2015719			81	1	1	6	
PRJNA283498	RNA-seq	SRR2015720			79	1	1	8.4	
PRJNA283498	RNA-seq	SRR2015721			63	1	2	6.5	
PRJNA283498	RNA-seq	SRR2015722			66	1		7.1	
PRJNA283498	RNA-seq	SRR2015723			69	1		7.8	
PRJNA283498	RNA-seq	SRR2015724			79	1		8	
PRJNA283498	RNA-seq	SRR2015725			61	1	1	8.2	
PRJNA283498	RNA-seq	SRR2015726			58	1	0	8.4	
PRJNA283498	RNA-seq	SRR2015727			70	1	1	8.2	
PRJNA283498	RNA-seq	SRR2015728			66	1	1	8.5	
PRJNA283498	RNA-seq	SRR2015729			73	1	1	7.5	
PRJNA283498	RNA-seq	SRR2015730			60	1		7.9	
PRJNA283498	RNA-seq	SRR2015731			76	1	1	7.3	
PRJNA283498	RNA-seq	SRR2015732			61	1	0	7.8	
PRJNA283498	RNA-seq	SRR2015733			62	1	1	6.6	
PRJNA283498	RNA-seq	SRR2015734			69	1	1	8.7	
PRJNA283498	RNA-seq	SRR2015735			61	1	1	8.1	
PRJNA283498	RNA-seq	SRR2015736			88	1	2	7.1	
PRJNA283498	RNA-seq	SRR2015737			93	1		6.4	
PRJNA283498	RNA-seq	SRR2015738			53	1	0	7.3	
PRJNA283498	RNA-seq	SRR2015739			57	1	1	8.3	
PRJNA283498	RNA-seq	SRR2015740			46	1	0	7.6	
PRJNA283498	RNA-seq	SRR2015741			57	1	1	7.7	
PRJNA283498	RNA-seq	SRR2015742			80	1	1	7.3	
PRJNA283498	RNA-seq	SRR2015743			74	1	2	8.5	
PRJNA283498	RNA-seq	SRR2015744			69	1	1	8.4	
PRJNA283498	RNA-seq	SRR2015745			76	1	1	7.5	
PRJNA283498	RNA-seq	SRR2015746			78	1	2	7.6	
PRJNA283498	RNA-seq	SRR2015747			87	1	4	8.7	
PRJNA283498	RNA-seq	SRR2015748			86	1	2	8.7	

PRJNA283498	RNA-seq	SRR2015749			54	1		8.3	
PRJNA283498	RNA-seq	SRR2015750			68	1	0	6.3	
PRJNA283498	RNA-seq	SRR2015751			52	1	0	7.4	
PRJNA283498	RNA-seq	SRR2015752			46	1	0	8.2	
PRJNA283498	RNA-seq	SRR2015753			55	1		7.6	
PRJNA283498	RNA-seq	SRR2015754			57	1	1	7.8	
PRJNA283498	RNA-seq	SRR2015755			66	1	1	8.4	
PRJNA283498	RNA-seq	SRR2015756			64	1	1	8.7	
PRJNA557205	RNA-seq	SRR9854867			79	2		7.5	
PRJNA557205	RNA-seq	SRR9854868			79	2		7.5	
PRJNA557205	RNA-seq	SRR9854869			79	2		7.5	
PRJNA557205	RNA-seq	SRR9854870			79	2		7.5	
PRJNA557205	RNA-seq	SRR9854871			82	1		8.3	
PRJNA557205	RNA-seq	SRR9854872			82	1		8.3	
PRJNA557205	RNA-seq	SRR9854873			82	1		8.3	
PRJNA557205	RNA-seq	SRR9854874			82	1		8.3	
PRJNA557205	RNA-seq	SRR9854875			70	1		7.7	
PRJNA557205	RNA-seq	SRR9854876			70	1		7.7	
PRJNA557205	RNA-seq	SRR9854877			70	1		7.7	
PRJNA557205	RNA-seq	SRR9854878			70	1		7.7	
PRJNA557205	RNA-seq	SRR9854879			77	1		7.1	
PRJNA557205	RNA-seq	SRR9854880			77	1		7.1	
PRJNA557205	RNA-seq	SRR9854881			77	1		7.1	
PRJNA557205	RNA-seq	SRR9854882			77	1		7.1	
PRJNA557205	RNA-seq	SRR9854883			73	1		7.8	
PRJNA557205	RNA-seq	SRR9854884			73	1		7.8	
PRJNA557205	RNA-seq	SRR9854885			73	1		7.8	
PRJNA557205	RNA-seq	SRR9854886			73	1		7.8	
PRJNA557205	RNA-seq	SRR9854887			73	1		7.1	
PRJNA557205	RNA-seq	SRR9854888			73	1		7.1	
PRJNA557205	RNA-seq	SRR9854889			73	1		7.1	
PRJNA557205	RNA-seq	SRR9854890			73	1		7.1	
PRJNA557205	RNA-seq	SRR9854891			68	1		8.1	
PRJNA557205	RNA-seq	SRR9854892			68	1		8.1	
PRJNA557205	RNA-seq	SRR9854893			68	1		8.1	
PRJNA557205	RNA-seq	SRR9854894			68	1		8.1	
PRJNA557205	RNA-seq	SRR9854895			68	1		7.6	
PRJNA557205	RNA-seq	SRR9854896			68	1		7.6	
PRJNA557205	RNA-seq	SRR9854897			68	1		7.6	
PRJNA557205	RNA-seq	SRR9854898			68	1		7.6	
PRJNA557205	RNA-seq	SRR9854899			84	1		8.2	

PRJNA557205	RNA-seq	SRR9854900	AD		84	1		8.2	
PRJNA557205	RNA-seq	SRR9854901			84	1		8.2	
PRJNA557205	RNA-seq	SRR9854902			84	1		8.2	
PRJNA557205	RNA-seq	SRR9854903			71	1		8.8	
PRJNA557205	RNA-seq	SRR9854904			71	1		8.8	
PRJNA557205	RNA-seq	SRR9854905			71	1		8.8	
PRJNA557205	RNA-seq	SRR9854906			71	1		8.8	
PRJNA557205	RNA-seq	SRR9854907			83	1		7.3	
PRJNA557205	RNA-seq	SRR9854908			83	1		7.3	
PRJNA557205	RNA-seq	SRR9854909			83	1		7.3	
PRJNA557205	RNA-seq	SRR9854910			83	1		7.3	
PRJNA557205	RNA-seq	SRR9854911			70	1		8	
PRJNA557205	RNA-seq	SRR9854912			70	1		8	
PRJNA557205	RNA-seq	SRR9854913			70	1		8	
PRJNA557205	RNA-seq	SRR9854914			70	1		8	
PRJNA143627	RNA-seq	SRR306841			91	1		8.5	
PRJNA185476	RNA-seq	SRR646457			54	2		7.3	
PRJNA185476	RNA-seq	SRR646458			67	2		6.1	
PRJNA185476	RNA-seq	SRR646459			56	2		7.7	
PRJNA185476	RNA-seq	SRR646460			44	2		6.8	
PRJEB5212	RNA-seq	ERR409878			43	1		6.8	
PRJEB5212	RNA-seq	ERR409879			80	1		8.2	
PRJEB5212	RNA-seq	ERR409882			53	1		7.6	
PRJEB5212	RNA-seq	ERR409883			46	1		7.8	
PRJEB5212	RNA-seq	ERR409887			76	1		8.1	
PRJEB5212	RNA-seq	ERR409888			57	1		8.1	
PRJEB5212	RNA-seq	ERR409891			79	1		7	
PRJEB5212	RNA-seq	ERR409893			40	1		7.4	
PRJEB5212	RNA-seq	ERR409896			57	1		8	
PRJEB5212	RNA-seq	ERR409898			93	1		7.8	
PRJEB5212	RNA-seq	ERR409900			44	1		8	
PRJNA678865	RNA-seq	SRR13071573	AD	PBMC	70	2			
PRJNA678865	RNA-seq	SRR13071574			73	2			
PRJNA678865	RNA-seq	SRR13071575			76	2			
PRJNA678865	RNA-seq	SRR13071576			76	2			
PRJNA678865	RNA-seq	SRR13071577			75	2			
PRJNA678865	RNA-seq	SRR13071585			70	2			
PRJNA678865	RNA-seq	SRR13071586			73	2			
PRJNA678865	RNA-seq	SRR13071587			76	2			
PRJNA678865	RNA-seq	SRR13071588			76	2			
PRJNA678865	RNA-seq	SRR13071589			75	2			

PRJNA667860	AMPLICON	SRR12786082			66	2			
PRJNA667860	AMPLICON	SRR12786083			76	1			
PRJNA667860	AMPLICON	SRR12786084			67	2			
PRJNA667860	AMPLICON	SRR12786085			67	2			
PRJNA667860	AMPLICON	SRR12786086			69	2			
PRJNA667860	AMPLICON	SRR12786087			68	2			
PRJNA667860	AMPLICON	SRR12786088			68	2			
PRJNA667860	AMPLICON	SRR12786089			81	2			
PRJNA667860	AMPLICON	SRR12786094			63	1			
PRJNA667860	AMPLICON	SRR12786095			78	2			
PRJNA667860	AMPLICON	SRR12786096			79	2			
PRJNA667860	AMPLICON	SRR12786097			61	1			
PRJNA678865	RNA-seq	SRR13071578	Ctrl	PBMC	76	2			
PRJNA678865	RNA-seq	SRR13071579			76	2			
PRJNA678865	RNA-seq	SRR13071584			74	1			
PRJNA678865	RNA-seq	SRR13071590			76	2			
PRJNA678865	RNA-seq	SRR13071591			76	2			
PRJNA678865	RNA-seq	SRR13071592			74	1			
PRJNA678865	RNA-seq	SRR13071593			61	1			
PRJNA678865	RNA-seq	SRR13071594			58	1			
PRJNA678865	RNA-seq	SRR13071595			61	1			
PRJNA678865	RNA-seq	SRR13071596			58	1			
PRJNA667860	AMPLICON	SRR12786090			72	1			
PRJNA667860	AMPLICON	SRR12786091			78	1			
PRJNA667860	AMPLICON	SRR12786092			78	1			
PRJNA667860	AMPLICON	SRR12786093			76	1			
PRJNA675864	RNA-seq	SRR13019398	PD	PBMC	62	1			
PRJNA675864	RNA-seq	SRR13019399			68	1			
PRJNA675864	RNA-seq	SRR13019400			56	1			
PRJNA675864	RNA-seq	SRR13019401			67	1			
PRJNA675864	RNA-seq	SRR13019402			64	2			
PRJNA675864	RNA-seq	SRR13019403			58	2			
PRJNA799936	RNA-seq	SRR17717617			50	1			
PRJNA799936	RNA-seq	SRR17717618			50	1			
PRJNA799936	RNA-seq	SRR17717620			68	1			
PRJNA799936	RNA-seq	SRR17717621			68	1			
PRJNA799936	RNA-seq	SRR17717622			68	1			
PRJNA799936	RNA-seq	SRR17717623			68	1			
PRJNA799936	RNA-seq	SRR17717624			69	1			
PRJNA799936	RNA-seq	SRR17717625			69	1			
PRJNA799936	RNA-seq	SRR17717626			69	1			

PRJNA799936	RNA-seq	SRR17717627			69	1			
PRJNA799936	RNA-seq	SRR17717628			68	2			
PRJNA799936	RNA-seq	SRR17717629			68	2			
PRJNA799936	RNA-seq	SRR17717630			68	2			
PRJNA799936	RNA-seq	SRR17717631			68	2			
PRJNA799936	RNA-seq	SRR17717632			56	2			
PRJNA799936	RNA-seq	SRR17717633			56	2			
PRJNA799936	RNA-seq	SRR17717634			56	2			
PRJNA799936	RNA-seq	SRR17717635			56	2			
PRJNA799936	RNA-seq	SRR17717637			50	1			
PRJNA799936	RNA-seq	SRR17717638			50	1			
PRJNA799936	RNA-seq	SRR17717640			65	1			
PRJNA799936	RNA-seq	SRR17717641			65	1			
PRJNA799936	RNA-seq	SRR17717645			65	1			
PRJNA799936	RNA-seq	SRR17717646			65	1			
PRJNA799936	RNA-seq	SRR17717647			50	1			
PRJNA799936	RNA-seq	SRR17717648			61	2			
PRJNA799936	RNA-seq	SRR17717649			61	2			
PRJNA799936	RNA-seq	SRR17717650			68	1			
PRJNA799936	RNA-seq	SRR17717651			69	1			
PRJNA799936	RNA-seq	SRR17717652			65	1			
PRJNA799936	RNA-seq	SRR17717653			63	2			
PRJNA799936	RNA-seq	SRR17717654			63	2			
PRJNA799936	RNA-seq	SRR17717655			61	2			
PRJNA799936	RNA-seq	SRR17717656			63	2			
PRJNA799936	RNA-seq	SRR17717657			63	2			
PRJNA799936	RNA-seq	SRR17717658			68	2			
PRJNA799936	RNA-seq	SRR17717659			56	2			
PRJNA799936	RNA-seq	SRR17717660			63	2			
PRJNA799936	RNA-seq	SRR17717661			61	2			
PRJNA799936	RNA-seq	SRR17717662			61	2			
PRJNA675864	RNA-seq	SRR13019394	Ctrl	PBMC	53	2			
PRJNA675864	RNA-seq	SRR13019395			61	1			
PRJNA675864	RNA-seq	SRR13019396			45	1			
PRJNA675864	RNA-seq	SRR13019397			56	1			
PRJNA675864	RNA-seq	SRR13019410			60	2			
PRJNA675864	RNA-seq	SRR13019411			57	1			
PRJNA675864	RNA-seq	SRR13019412			48	2			
PRJNA675864	RNA-seq	SRR13019413			59	1			
PRJNA675864	RNA-seq	SRR13019414			51	1			
PRJNA675864	RNA-seq	SRR13019415			44	1			

PRJNA675864	RNA-seq	SRR13019416	AD		50	1			
PRJNA799936	RNA-seq	SRR17717593			65	1			
PRJNA799936	RNA-seq	SRR17717594			65	1			
PRJNA799936	RNA-seq	SRR17717595			65	1			
PRJNA799936	RNA-seq	SRR17717596			65	1			
PRJNA799936	RNA-seq	SRR17717597			65	1			
PRJNA799936	RNA-seq	SRR17717598			65	1			
PRJNA799936	RNA-seq	SRR17717599			65	1			
PRJNA799936	RNA-seq	SRR17717600			65	1			
PRJNA799936	RNA-seq	SRR17717601			72	1			
PRJNA799936	RNA-seq	SRR17717602			72	1			
PRJNA799936	RNA-seq	SRR17717603			72	1			
PRJNA799936	RNA-seq	SRR17717604			72	1			
PRJNA799936	RNA-seq	SRR17717605			51	1			
PRJNA799936	RNA-seq	SRR17717606			51	1			
PRJNA799936	RNA-seq	SRR17717607			51	1			
PRJNA799936	RNA-seq	SRR17717608			51	1			
PRJNA799936	RNA-seq	SRR17717609			65	2			
PRJNA799936	RNA-seq	SRR17717610			65	2			
PRJNA799936	RNA-seq	SRR17717611			65	2			
PRJNA799936	RNA-seq	SRR17717612			65	2			
PRJNA799936	RNA-seq	SRR17717613			71	2			
PRJNA799936	RNA-seq	SRR17717614			71	2			
PRJNA799936	RNA-seq	SRR17717615			71	2			
PRJNA799936	RNA-seq	SRR17717616			71	2			
PRJNA799936	RNA-seq	SRR17717619			65	1			
PRJNA799936	RNA-seq	SRR17717636			65	1			
PRJNA799936	RNA-seq	SRR17717639			72	1			
PRJNA799936	RNA-seq	SRR17717642			51	1			
PRJNA799936	RNA-seq	SRR17717643			71	2			
PRJNA799936	RNA-seq	SRR17717644			65	2			
PRJNA201039	RNA-seq	SRR837437	AD	WB	77	2			
PRJNA201039	RNA-seq	SRR837438			74	1			
PRJNA201039	RNA-seq	SRR837439			68	1			
PRJNA201039	RNA-seq	SRR837440			75	2			
PRJNA201039	RNA-seq	SRR837441			74	1			
PRJNA201039	RNA-seq	SRR837442			76	1			
PRJNA201039	RNA-seq	SRR837443			79	1			
PRJNA201039	RNA-seq	SRR837444			75	2			
PRJNA201039	RNA-seq	SRR837445			77	1			
PRJNA201039	RNA-seq	SRR837446			72	2			

PRJNA201039	RNA-seq	SRR837447			76	1			
PRJNA201039	RNA-seq	SRR837448			61	1			
PRJNA201039	RNA-seq	SRR837449			51	2			
PRJNA201039	RNA-seq	SRR837450			55	2			
PRJNA201039	RNA-seq	SRR837451			70	2			
PRJNA201039	RNA-seq	SRR837452			78	2			
PRJNA201039	RNA-seq	SRR837453			81	1			
PRJNA201039	RNA-seq	SRR837454			72	2			
PRJNA201039	RNA-seq	SRR837455			57	1			
PRJNA201039	RNA-seq	SRR837456			75	1			
PRJNA201039	RNA-seq	SRR837457			70	2			
PRJNA201039	RNA-seq	SRR837458			75	2			
PRJNA201039	RNA-seq	SRR837459			69	2			
PRJNA201039	RNA-seq	SRR837460			68	2			
PRJNA201039	RNA-seq	SRR837461			71	2			
PRJNA201039	RNA-seq	SRR837462			74	1			
PRJNA201039	RNA-seq	SRR837463			72	1			
PRJNA201039	RNA-seq	SRR837464			69	2			
PRJNA201039	RNA-seq	SRR837465			63	2			
PRJNA201039	RNA-seq	SRR837466			80	1			
PRJNA201039	RNA-seq	SRR837467			62	1			
PRJNA201039	RNA-seq	SRR837468			75	1			
PRJNA201039	RNA-seq	SRR837469			56	2			
PRJNA201039	RNA-seq	SRR837470			57	2			
PRJNA201039	RNA-seq	SRR837493			80	1			
PRJNA201039	RNA-seq	SRR837494			81	1			
PRJNA201039	RNA-seq	SRR837495			76	2			
PRJNA201039	RNA-seq	SRR837496			60	2			
PRJNA201039	RNA-seq	SRR837497			63	2			
PRJNA201039	RNA-seq	SRR837498			64	2			
PRJNA201039	RNA-seq	SRR837499			65	1			
PRJNA201039	RNA-seq	SRR837500			74	2			
PRJNA201039	RNA-seq	SRR837501			59	2			
PRJNA201039	RNA-seq	SRR837502			72	2			
PRJNA201039	RNA-seq	SRR837503			63	1			
PRJNA201039	RNA-seq	SRR837504			78	1			
PRJNA201039	RNA-seq	SRR837505			78	1			
PRJNA201039	RNA-seq	SRR837506			79	1			
PRJNA714081	RNA-seq	SRR13953402				2			
PRJNA714081	RNA-seq	SRR13953405				2			
PRJNA714081	RNA-seq	SRR13953401				2			

PRJNA714081	RNA-seq	SRR13953403				2			
PRJNA714081	RNA-seq	SRR13953404				2			
PRJNA201039	RNA-seq	SRR837471	Ctrl	WB	76	1			
PRJNA201039	RNA-seq	SRR837472			60	2			
PRJNA201039	RNA-seq	SRR837473			60	2			
PRJNA201039	RNA-seq	SRR837474			76	2			
PRJNA201039	RNA-seq	SRR837475			60	2			
PRJNA201039	RNA-seq	SRR837476			83	1			
PRJNA201039	RNA-seq	SRR837477			61	1			
PRJNA201039	RNA-seq	SRR837478			77	1			
PRJNA201039	RNA-seq	SRR837479			63	1			
PRJNA201039	RNA-seq	SRR837480			75	1			
PRJNA201039	RNA-seq	SRR837481			63	1			
PRJNA201039	RNA-seq	SRR837482			64	1			
PRJNA201039	RNA-seq	SRR837483			71	2			
PRJNA201039	RNA-seq	SRR837484			60	2			
PRJNA201039	RNA-seq	SRR837485			75	1			
PRJNA201039	RNA-seq	SRR837486			60	2			
PRJNA201039	RNA-seq	SRR837487			66	2			
PRJNA201039	RNA-seq	SRR837488			60	2			
PRJNA201039	RNA-seq	SRR837489			61	1			
PRJNA201039	RNA-seq	SRR837490			74	2			
PRJNA201039	RNA-seq	SRR837491			71	2			
PRJNA201039	RNA-seq	SRR837492			60	1			
PRJNA714081	RNA-seq	SRR13953391				2			
PRJNA714081	RNA-seq	SRR13953393				2			
PRJNA714081	RNA-seq	SRR13953394				2			
PRJNA714081	RNA-seq	SRR13953397				2			
PRJNA714081	RNA-seq	SRR13953398				2			
PRJNA714081	RNA-seq	SRR13953400				2			
PRJNA714081	RNA-seq	SRR13953392				2			
PRJNA714081	RNA-seq	SRR13953395				2			
PRJNA714081	RNA-seq	SRR13953396				2			
PRJNA714081	RNA-seq	SRR13953399				2			
PRJEB8960	RNA-seq	ERR833219	PD	WB	59	2			
PRJEB8960	RNA-seq	ERR833220			59	2			
PRJEB8960	RNA-seq	ERR833221			59	2			
PRJEB8960	RNA-seq	ERR833222			59	2			
PRJEB8960	RNA-seq	ERR833227			53	1			
PRJEB8960	RNA-seq	ERR833228			53	1			
PRJEB8960	RNA-seq	ERR833229			53	1			

PRJEB8960	RNA-seq	ERR833230			53	1			
PRJEB8960	RNA-seq	ERR833239			50	1			
PRJEB8960	RNA-seq	ERR833240			50	1			
PRJEB8960	RNA-seq	ERR833241			50	1			
PRJEB8960	RNA-seq	ERR833242			50	1			
PRJEB8960	RNA-seq	ERR833251			47	1			
PRJEB8960	RNA-seq	ERR833252			47	1			
PRJEB8960	RNA-seq	ERR833253			47	1			
PRJEB8960	RNA-seq	ERR833254			47	1			
PRJEB8960	RNA-seq	ERR833255			50	1			
PRJEB8960	RNA-seq	ERR833256			50	1			
PRJEB8960	RNA-seq	ERR833257			50	1			
PRJEB8960	RNA-seq	ERR833258			50	1			
PRJNA675864	RNA-seq	SRR13019398			62	1			
PRJNA675864	RNA-seq	SRR13019399			68	1			
PRJNA675864	RNA-seq	SRR13019400			56	1			
PRJNA675864	RNA-seq	SRR13019401			67	1			
PRJNA675864	RNA-seq	SRR13019402			64	2			
PRJNA675864	RNA-seq	SRR13019403			58	2			
PRJNA693202	RNA-seq	SRR13482599			71	2			
PRJNA693202	RNA-seq	SRR13482600			78	1			
PRJNA693202	RNA-seq	SRR13482601			77	2			
PRJNA693202	RNA-seq	SRR13482602			72	1			
PRJNA693202	RNA-seq	SRR13482603			75	2			
PRJNA693202	RNA-seq	SRR13482604			75	2			
PRJNA693202	RNA-seq	SRR13482605			76	1			
PRJNA693202	RNA-seq	SRR13482606			74	2			
PRJNA693202	RNA-seq	SRR13482607			77	1			
PRJNA693202	RNA-seq	SRR13482608			78	1			
PRJNA693202	RNA-seq	SRR13482609			71	2			
PRJNA693202	RNA-seq	SRR13482610			72	1			
PRJEB8960	RNA-seq	ERR833223	Ctrl	WB	54	2			
PRJEB8960	RNA-seq	ERR833224			54	2			
PRJEB8960	RNA-seq	ERR833225			54	2			
PRJEB8960	RNA-seq	ERR833226			54	2			
PRJEB8960	RNA-seq	ERR833231			45	1			
PRJEB8960	RNA-seq	ERR833232			45	1			
PRJEB8960	RNA-seq	ERR833233			45	1			
PRJEB8960	RNA-seq	ERR833234			45	1			
PRJEB8960	RNA-seq	ERR833235			57	1			
PRJEB8960	RNA-seq	ERR833236			57	1			

PRJEB8960	RNA-seq	ERR833237			57	1			
PRJEB8960	RNA-seq	ERR833238			57	1			
PRJEB8960	RNA-seq	ERR833243			42	2			
PRJEB8960	RNA-seq	ERR833244			42	2			
PRJEB8960	RNA-seq	ERR833245			42	2			
PRJEB8960	RNA-seq	ERR833246			42	2			
PRJEB8960	RNA-seq	ERR833247			42	1			
PRJEB8960	RNA-seq	ERR833248			42	1			
PRJEB8960	RNA-seq	ERR833249			42	1			
PRJEB8960	RNA-seq	ERR833250			42	1			
PRJNA675864	RNA-seq	SRR13019394			53	2			
PRJNA675864	RNA-seq	SRR13019395			61	1			
PRJNA675864	RNA-seq	SRR13019396			45	1			
PRJNA675864	RNA-seq	SRR13019397			56	1			
PRJNA675864	RNA-seq	SRR13019410			60	2			
PRJNA675864	RNA-seq	SRR13019411			57	1			
PRJNA675864	RNA-seq	SRR13019412			48	2			
PRJNA675864	RNA-seq	SRR13019413			59	1			
PRJNA675864	RNA-seq	SRR13019414			51	1			
PRJNA675864	RNA-seq	SRR13019415			44	1			
PRJNA675864	RNA-seq	SRR13019416			50	1			
PRJNA693202	RNA-seq	SRR13482611			68	2			
PRJNA693202	RNA-seq	SRR13482612			70	1			
PRJNA693202	RNA-seq	SRR13482613			73	2			
PRJNA693202	RNA-seq	SRR13482614			69	2			
PRJNA693202	RNA-seq	SRR13482615			72	2			
PRJNA693202	RNA-seq	SRR13482616			73	1			
PRJNA693202	RNA-seq	SRR13482617			70	2			
PRJNA693202	RNA-seq	SRR13482618			68	1			
PRJNA693202	RNA-seq	SRR13482619			68	2			
PRJNA693202	RNA-seq	SRR13482620			70	1			
PRJNA693202	RNA-seq	SRR13482621			72	2			
PRJNA693202	RNA-seq	SRR13482622			73	1			
PRJNA693202	RNA-seq	SRR13482623			69	2			
PRJNA693202	RNA-seq	SRR13482624			71	1			